nature biotechnology

An improved zinc-finger nuclease architecture for highly specific genome editing

Jeffrey C Miller[1], Michael C Holmes[1], Jianbin Wang[1], Dmitry Y Guschin[1], Ya-Li Lee[1], Igor Rupniewski[1], Christian M Beausejour[1,2], Adam J Waite[1], Nathaniel S Wang[1], Kenneth A Kim[1], Philip D Gregory[1], Carl O Pabo[1,2] & Edward J Rebar[1]

Genome editing driven by zinc-finger nucleases (ZFNs) yields high gene-modification efficiencies (>10%) by introducing a recombinogenic double-strand break into the targeted gene. The cleavage event is induced using two custom-designed ZFNs that heterodimerize upon binding DNA to form a catalytically active nuclease complex. Using the current ZFN architecture, however, cleavage-competent homodimers may also form that can limit safety or efficacy via off-target cleavage. Here we develop an improved ZFN architecture that eliminates this problem. Using structure-based design, we engineer two variant ZFNs that efficiently cleave DNA only when paired as a heterodimer. These ZFNs modify a native endogenous locus as efficiently as the parental architecture, but with a >40-fold reduction in homodimer function and much lower levels of genome-wide cleavage. This architecture provides a general means for improving the specificity of ZFNs as gene modification reagents.

Zinc-finger nucleases (ZFNs) are rapidly emerging as versatile reagents for gene modification. These hybrid restriction enzymes, which link the cleavage domain of *Fok*I to a designed zinc-finger protein (ZFP)[1,2], may be used to introduce a variety of custom alterations into the genomes of eukaryotic cells. Examples range from precise sequence edits[3–6] to the targeted integration of entire genes[7]. ZFNs initiate these events by introducing a double-strand break at the site chosen for modification. If an exogenous repair template is also supplied then sequence alterations encoded in this donor may be incorporated into the genome by homology-directed repair (HDR)[8]. Targeting of the cleavage event, which is central to ZFN specificity and versatility, is mediated by the ZFP domain. This DNA-binding domain has been characterized in great detail[9–11] and may be engineered to recognize a wide variety of chosen DNA sequences[12–16].

ZFNs may offer a general method for engineering the genomes of diverse species as the requisite DNA repair pathways are highly conserved[17]. ZFN-stimulated gene modification has been demonstrated in plants[18,19], insects[4,6,20], roundworms[21] and human cells[3,5], with endogenous gene correction rates as high as 18%[5]. Moreover, donor-free delivery of ZFNs may be used for the targeted disruption of endogenous genes[4,6,20]. The generality of ZFN-mediated gene modification raises the prospect of applications in diverse fields including crop engineering, therapeutic gene correction, cell line customization for biologics production and the development of nonmurine models of human disease. Realizing the full potential of these approaches, however, will require the development of ZFN architectures and design strategies that yield efficient cleavage while minimizing off-target effects.

An important feature of the current ZFN architecture is that DNA cleavage requires nuclease dimerization. ZFNs, as well as wild-type *Fok*I, interact through their cleavage domains and are inactive as monomers[22–25]. As a consequence, cleavage of a typical target requires the design of two different ZFNs for binding to adjacent half-sites (**Fig. 1a**,**b**). Moreover, the dimerization interaction is quite weak; *Fok*I remains monomeric at concentrations of at least 15 μM and dimerizes only when bound to its specific target[23,26]. From the standpoint of targeting specificity, an advantage of this arrangement is that cleavage requires simultaneous binding of both ZFNs to their respective half-sites. This has enabled the construction of ZFN dimers with cleavage specificities of up to 24 bp[5], a length that exceeds the target size of other proposed gene modification agents such as the meganuclease I-*Sce* I[27] and that offers the prospect for unique targeting within the human genome. The development of ZFNs that specifically cleave even longer targets may be anticipated, as individual ZFPs have been successfully designed for sequences as long as 18 bp[28–31].

Although the requirement for dimerization opens up the possibility of restricting cleavage to very long and rare sequences, it also introduces a problem arising from the fact that protein-protein interactions mediated by the wild-type *Fok*I cleavage domain are not themselves selective for the heterodimer species (**Fig. 1c**). As a consequence, the expression of any ZFN heterodimer (e.g., L^{wt}/R^{wt} in **Fig. 1c**) also yields two side-product homodimers (L^{wt}/L^{wt} and R^{wt}/R^{wt}) that will be irrelevant for gene modification but may nonetheless limit safety or effectiveness by cleaving off-target sequences. Moreover, if one ZFN is less specific than the other, the

[1]Sangamo BioSciences, Inc., Pt. Richmond Tech Center, 501 Canal Blvd., Suite A100 Richmond, California 94804, USA. [2]Present addresses: Département de pharmacologie, Centre de Recherche, CHU Ste-Justine 3175, Côte Ste-Catherine, Montréal, Quebec H3T 1C5, Canada (C.M.B.) and Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, WAB 536, Boston, Massachusetts 02115, USA (C.O.P.). Correspondence and requests for materials should be addressed to E.J.R. (erebar@sangamo.com).

Received 31 January; accepted 4 June; published online 1 July 2007; doi:10.1038/nbt1319

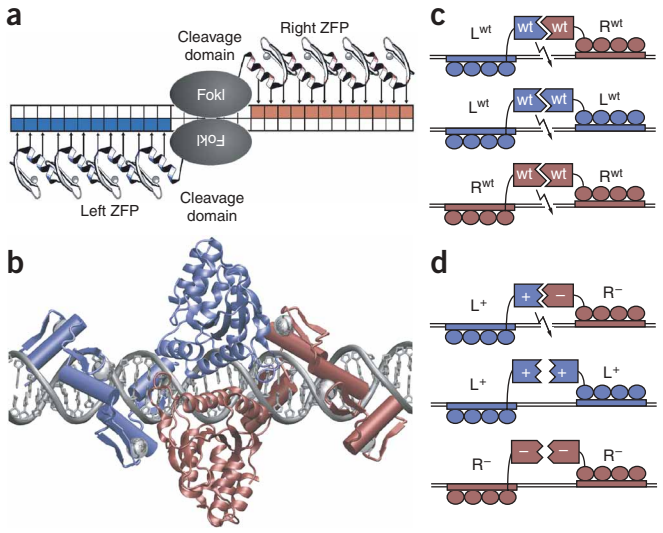

Figure 1 DNA recognition and cleavage by zinc-finger nucleases (ZFNs). (**a**) Sketch of a ZFN dimer bound to a typical, nonpalindromic DNA target. Each ZFN consists of the cleavage domain of *Fok*I fused to a zinc-finger protein (ZFP) that has been customized to specifically recognize either a 'left' or 'right' half-site (indicated by blue and red boxes), which are separated by a spacer of either 5 or 6 bp. Simultaneous binding by both ZFNs enables dimerization of the *Fok*I nuclease domain and DNA cleavage. Note that endogenously active ZFNs have been created using four fingers (indicated here) that each bind 12-bp sites[5], as well as three fingers that each bind 9-bp sites[4,6,20,21]. (**b**) Three-dimensional model corresponding to the arrangement in **a**. DNA is shown in gray, ZFNs are colored blue (left ZFN) or red (right ZFN), and reflective spheres denote zinc ions. Zinc-finger helices that mediate sequence recognition in the major groove are represented by cylinders, whereas the more centrally located *Fok*I cleavage domains are represented entirely as ribbon diagrams. The cleavage domain interface that has been reengineered in these studies is located at the center of the image. (**c**) Sketch of the problem motivating this work. Coexpression of L^{wt} (left ZFP fused to the wild-type *Fok*I cleavage domain) and R^{wt} (right ZFP fused to wild-type *Fok*I cleavage domain) yields a heterodimer which cleaves the desired target (top) but also yields homodimers that may cleave other targets (middle and bottom). Wild-type *Fok*I cleavage domains are indicated by self-complementary polygons labeled 'wt'. ZFP domains are indicated by four adjacent circles. Jagged arrows highlight the capacity of each dimer species to cleave DNA. (**d**) Sketch of the solution offered by the new protein designs. Variants of the *Fok*I cleavage domain that function as obligate heterodimers (complementary polygons labeled '+' and '–') would enable cleavage of desired heterodimer targets (top) while eliminating off-target activity by homodimer species (middle and bottom). The labels L^+ and $R^–$ denote, respectively, the 'left' and 'right' ZFPs fused to the variant cleavage domains. The absence of a jagged arrow with the L^+/L^+ and $R^–/R^–$ pairings denotes the inability of these homodimers to cleave DNA. Figures were made using VMD (http://www.ks.uiuc.edu/Research/vmd/) and rendered using POV-Ray (http://www.povray.org/).

homodimer of the less specific ZFN will tend to be the dose-limiting species in terms of off-target effects. Evidence of this phenomenon was seen in a recent study, in which two out of three designed ZFN heterodimers exhibited substantial toxicity that limited effectiveness and that was, in each case, due to homodimer formation of a single ZFN[6]. Moreover, mutating the active site in one such ZFN relieved its toxicity, indicating that homodimer cleavage properties, rather than DNA binding, mediated the toxic effect[6].

The creation of *Fok*I cleavage domain variants that preferentially heterodimerize would provide a general solution to the problem of homodimer formation while preserving the advantages of dimerization-dependent cleavage (**Fig. 1d**). Here, we describe the development of such cleavage domain variants using a strategy of iterative structure-based design followed by screening for HDR-driven gene correction in mammalian cells. These variants, which function as obligate heterodimers, should be generally useful for improving the potency and specificity of ZFNs as gene-targeting reagents.

RESULTS

Experimental strategy

A variety of strategies have been described for constructing heterodimeric protein interaction surfaces[32–37]. The development of *Fok*I heterodimers, however, presented a novel combination of challenges that required a distinct approach. One critical difference from prior studies was that the target for redesign efforts, the *Fok*I cleavage domain, is an enzyme. It was therefore important to screen variant designs for catalytic function instead of just interaction preference. This was a concern because active-site residues are within 10 Å of the interface[38]. A second feature of these studies was the very low affinity of the *Fok*I dimer interaction. Our interest in preserving this property, which is important for ensuring cleavage specificity, argued against using affinity-based selection methods, such as phage display, for heterodimer development. A final consideration was the predominantly hydrophilic nature of the *Fok*I dimer interface, which contains numerous electrostatic interactions as well as many ordered waters[38]. This feature made it less likely that a purely computational strategy would yield monomers with the desired properties (although a recent study has reported substantial progress in this area[39]).

Given these considerations, our heterodimer development strategy involved two key elements. First, to ensure that variants retained the desired catalytic activity, we used functional screens that directly tested design candidates for the ability to cleave DNA. Our primary test platform involved an assay for green fluorescent protein (GFP) reporter gene correction in mammalian cells[3,5]. This system utilizes an integrated, constitutively expressed eGFP gene that is disrupted by insertion of the target for a well-characterized ZFN dimer[5] (**Fig. 2a,b**). The ZFN target sequence in our GFP reporter system is the binding site for the dimer ZFN-L*/ZFN-R*[5], which was designed for gene correction of the endogenous human IL2Rγ locus. For the sake of clarity, we will refer to the DNA-binding domains of these ZFNs as the 'L' and 'R' ZFPs throughout the rest of this paper.

The second component of our development strategy consisted of a stepwise approach to modification of the dimer interface. This was implemented through four cycles of variant design and testing, each of which substituted one amino acid at the dimer interface and yielded a progressive improvement in the specificity of heterodimer formation. In each cycle of development, a small panel of amino acid substitutions was generated within one cleavage domain, whereas its partner was not modified. The choice of substitutions was guided by the coordinates of the native *Fok*I dimer[38] (**Fig. 2c**). Mutations were made at positions that could contact the unmodified partner, with a bias toward the introduction of charge-charge interactions (**Supplementary Methods** and **Supplementary Table 1** online). Each variant was then screened for the ability to stimulate gene correction in two alternative configurations: (i) as a heterodimer with the unmodified partner ZFN, and (ii) as a homodimer. Successive development cycles alternated between the two sides of the dimer interface.

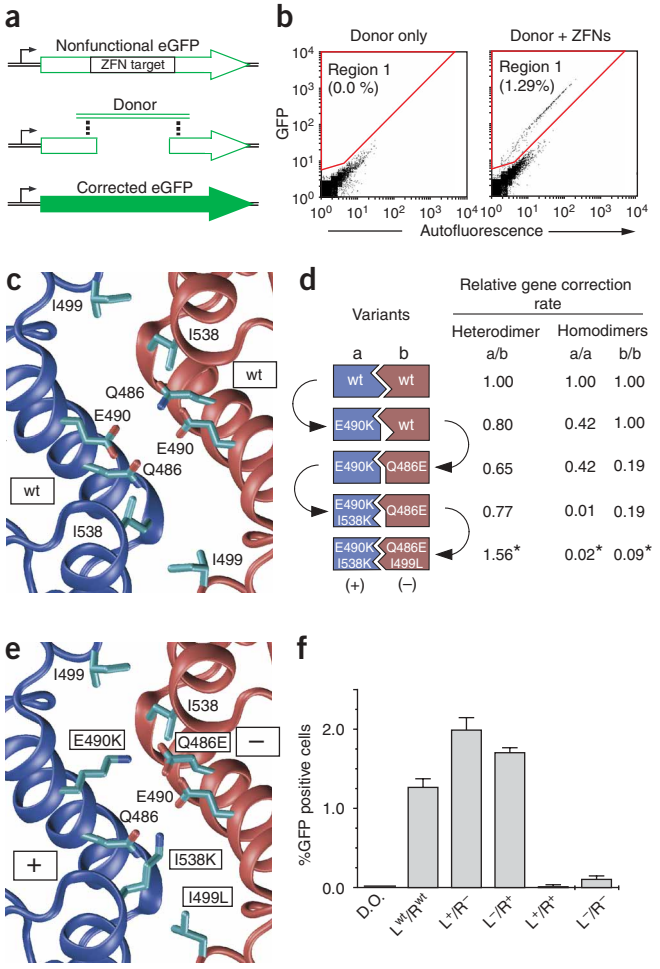

Figure 2 Development of *Fok*I cleavage domain mutants that function as obligate heterodimers. (**a**) Sketch of the reporter system used to screen cleavage domain variants for gene correction activity[3,5]. The system uses an HEK293 cell line that contains an integrated copy of the eGFP gene interrupted by a short DNA fragment bearing the cleavage target for a well-characterized ZFN dimer[5] (top). Cleavage and resection yields a substrate for homology-directed repair (HDR) which may use an exogenous donor DNA fragment bearing a portion of the eGFP sequence as a repair template (middle and bottom). Gene correction efficiency is monitored by counting the fraction of cells that convert to green fluorescence. The dashed line between the resected eGFP gene and the donor indicates the HDR process. (**b**) Representative flow cytometry data for reporter cells transfected with either the donor only or the donor plus ZFNs. The fraction of GFP-positive cells may be quantified using the indicated gate. (**c**) Central region of the dimer interface of wild-type *Fok*I. Labels highlight residues mutated in these studies. Coordinates are from reference 38. (**d**) Summary of the iterative screening process used to identify *Fok*I cleavage domain variants that function as obligate heterodimers. 'Wt' indicates the wild-type *Fok*I cleavage domain. *Fok*I cleavage domain variants are indicated by mutations relative to the wild-type sequence using the single-letter amino acid code. In each development step, curved arrows connect the parental cleavage domain to the discovered variant having the best heterodimerization properties. The arrows also highlight our strategy of alternately modifying each surface of the dimer interface in a stepwise fashion. This type of approach shares some features with the 'computational second-site suppressor' method described by Kortemme and coworkers[57], although here we break the symmetry of a homodimer interface by testing of a small panel of single amino acid substitutions and then use experimental activity data rather than binding energy calculations to guide subsequent modifications. The relative gene correction rates indicate the fraction of GFP-positive cells observed for the indicated complex normalized to the fraction of GFP-positive cells obtained with the wt/wt dimer. The % gene correction achieved with the wt/wt dimer was 1.27 ± 0.11 in the experiment yielding data highlighted with an asterisk and 0.93 ± 0.09 in the experiment yielding the remaining (nonhighlighted) data. Errors indicate s.d. (**e**) Model of the interface between the '+' and '–' cleavage domain heterodimer variants. Introduced mutations are boxed. Side chains are positioned to demonstrate possible interactions between the positively charged Lys490 and Lys538 side chains on the '+' variant with the negatively charged Glu486 and Glu490 side chains on the '–' variant, as well as a possible van der Waals contact from L499 to K538. These mutated residues may also mediate repulsive interactions or disrupt hydrophobic packing in a dimer containing two copies of the same domain. (**f**) Gene correction efficiencies of the '+' and '–'cleavage domain variants in the GFP reporter system. L^{wt}, 'left' ZFP fused to a wild-type *Fok*I cleavage domain (identical to ZFN-L* from reference 5); L^{+} and L^{-}, 'left' ZFP fused to '+' and '-' cleavage domain variants, respectively; R^{wt}, 'right' ZFP fused to a wild-type *Fok*I cleavage domain[5]; R^{+} and R^{-} denote the 'right' ZFP fused to the '+' and '–' cleavage domain variants; D.O., cells transfected with donor plasmid only. The variants stimulate efficient gene correction as heterodimers (compare L^{+}/R^{-} and L^{-}/R^{+} versus L^{wt}/R^{wt}) but not when forced to self-associate (L^{+}/R^{+} and L^{-}/R^{-}). Values represent the average of four independent transfections. Error bars denote s.d.

Development of heterodimeric *Fok*I cleavage domain variants

In each cycle of development, we were able to identify a variant that efficiently induced gene correction as a heterodimer with the constant partner ZFN but that exhibited reduced activity as a homodimer (**Fig. 2d**). The first development cycle identified a variant with a glutamic acid → lysine mutation at position 490 (E490K) that exhibited a modest preference for heterodimerization with the wild-type *Fok*I cleavage domain. This mutant showed 80% of parental *Fok*I activity when tested as a heterodimer with the wild-type cleavage domain but exhibited reduced activity (42% of parental *Fok*I) as a homodimer. In the second cycle (now focusing on the other side of the interface), candidate partners were tested for preferential association with the E490K mutant, and a variant (Q486E) was identified that had suitable properties. This variant showed 65% of parental *Fok*I activity when tested as a heterodimer with the E490K domain but exhibited reduced activity (19% of parental *Fok*I) as a homodimer.

Two subsequent development cycles yielded variant cleavage domains with the double mutations E490K:I538K and Q486E:I499L (**Fig. 2d**). For the sake of clarity, these mutant domains will be referred to as '+' and '–', respectively, throughout the remainder of this paper, with the notation reflecting the charges of the new residues. A model of how these variants may interact at the *Fok*I dimer interface is provided in **Fig. 2e**. These variants exhibit a strong preference for heterodimerization in the GFP reporter system (**Fig. 2f**). When tested for potentiation of gene correction, fusion of the '+' cleavage variant to the 'L' ZFP and of the '–' variant to the 'R' ZFP yielded a complex that stimulated gene correction efficiencies of 1.98% (**Fig. 2f**). Moreover, reversal of the ZFP–cleavage domain pairings (the L^{-}/R^{+} dimer) yielded a similarly high correction rate (1.71%). In contrast, when forced to homodimerize, both cleavage domain variants were strongly impaired, yielding correction rates of just 0.02% and 0.12% for the L^{+}/R^{+} and L^{-}/R^{-} homodimers, respectively. For comparison, the wild-type *Fok*I cleavage domain yielded a gene correction rate of 1.27% (**Fig. 2f**). Taken together, these GFP reporter results indicated that the '+' and '–' variants form an effective, obligate heterodimer pair.

Efficient endogenous gene editing by cleavage domain variants

Our use of the 'L' and 'R' ZFPs for these studies, which target a sequence present in exon 5 of the human IL2Rγ gene, enabled us next to confirm the function of our dimerization mutants at this

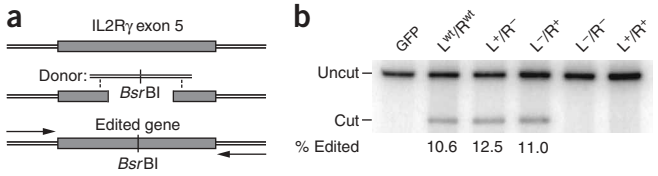

Figure 3 Gene editing activity of *Fok*I cleavage domain variants at a native endogenous target. (**a**) Sketch of endogenous gene-editing assay. Cleavage and resection of the native IL2Rγ gene (top and middle) yields a substrate for HDR, which may use a donor DNA fragment bearing a restriction site for *Bsr*BI as a repair template (middle and bottom). Gene editing efficiency is then quantified by PCR-amplification of the endogenous locus using primers that bind outside of the donor sequence followed by digestion of the PCR product with *Bsr*BI. The dashed line between the resected IL2Rγ gene and donor indicates the HDR process. (**b**) Results of the endogenous gene editing assay using human K562 cells transfected with the indicated constructs. 'GFP' indicates pmaxGFP (Amaxa), whereas other construct identities are as described in **Figure 2f**. 'Uncut' indicates the position of undigested PCR product and 'cut' indicates the position of the fragments created by *Bsr*BI digestion (which appear only if the *Bsr*BI restriction site has been transferred from the donor to the gene). The numbers below the lanes with visible *Bsr*BI bands indicate the percentage of cleaved PCR product.

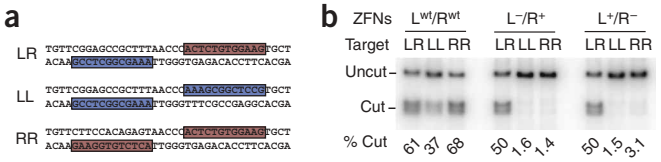

Figure 4 Preferential heterodimer activity of *Fok*I variants as demonstrated by DNA cleavage *in vitro*. (**a**) DNA target fragments. Both strands of DNA are shown with the top strand written 5′ to 3′ and the bottom strand written 3′ to 5′. Colors highlight the primary strand of the 12-bp sites for the 'L' (blue) and 'R' (red) ZFNs. (**b**) *In vitro* target cleavage by the indicated ZFN combinations. Migration of cleaved and uncleaved products is indicated. Numbers below each lane indicate the percentage of DNA in cleaved bands. The protein concentration of the L^{wt}/R^{wt} combination was adjusted to yield levels of LR target cleavage similar to that of the L^{+}/R^{-} and L^{-}/R^{+} combinations. Whereas homodimer formation can occur with the wild-type nuclease domain and enables cleavage of all three targets by the L^{wt}/R^{wt} combination (left panel), substantially reduced cleavage of the LL and RR targets is exhibited by ZFNs bearing the heterodimer cleavage domain variants (center and right panels).

endogenous locus. We assayed our ZFNs by using them to induce a gene editing event that introduces a novel *Bsr*BI restriction site[5] (**Fig. 3a**). The efficiency of gene editing was then quantified by PCR-amplification followed by *Bsr*BI digestion and gel electrophoresis (**Fig. 3b**). These studies showed that both the L^{+}/R^{-} and L^{-}/R^{+} ZFN combinations were able to induce gene editing at levels that were both highly efficient (>10%) and equivalent to those induced by the L^{wt}/R^{wt} ZFN combination. Importantly, forced homodimerization (L^{-}/R^{-} and L^{+}/R^{+}) yielded no detectable gene editing in this assay (**Fig. 3b**). These studies confirmed that the '+' and '–' variants behave as obligate heterodimers at a native endogenous gene target.

Suppression of homodimer function *in vitro*

Taken together, the endogenous gene editing studies and GFP reporter studies provided strong evidence that the '+' and '–' variants would substantially suppress ZFN homodimer function. To confirm this, we directly measured cleavage activities *in vitro*. Three radiolabeled duplexes were generated that contained target sequences for either the L/R heterodimer or corresponding L/L and R/R homodimers (**Fig. 4a**). Each target was then digested by the L and R ZFNs bearing either the wild-type *Fok*I cleavage domain (L^{wt}/R^{wt}) or the '+' and '–' variants in both possible configurations (L^{-}/R^{+} and L^{+}/R^{-}). We observed efficient cleavage of the heterodimer target by our cleavage domain variants with substantially less activity against the homodimer targets (**Fig. 4b**) and no activity when the same *Fok*I variant was used on both nucleases (**Supplementary Fig. 1** online). In contrast, a concentration of the L^{wt}/R^{wt} dimer yielding similar levels of activity against the heterodimer target also efficiently cleaved both the LL and RR targets owing to homodimer formation (**Fig. 4b**). These results confirmed the ability of the new variants to suppress ZFN homodimer function.

Reduced levels of genome-wide cleavage

Because the primary motivation for this work was to increase ZFN specificity in the context of gene-modification procedures, we next sought to determine directly whether the heterodimer variants reduced genome-wide DNA cleavage levels when expressed in mammalian cells. To accomplish this, we used two well-validated assays for visualizing DNA double-strand breaks that involve antibody-mediated detection of proteins associated with sites of DNA damage. In these studies, target cells were transfected with ZFN expression constructs alone (without donor DNA). Our initial studies used detection of the protein 53BP1, which localizes to sites of DNA damage and forms foci that may be visualized by

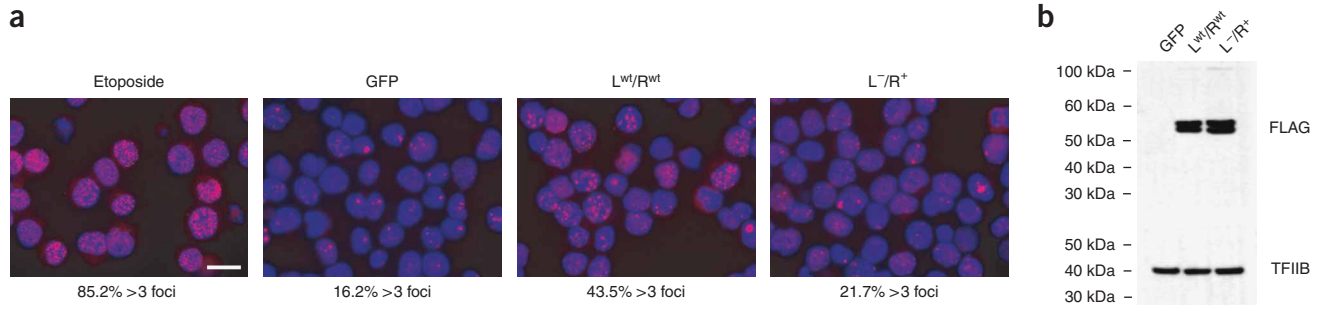

Figure 5 Reduced DNA damage levels by heterodimer variants. (**a**) Representative images of cells treated with the DNA-damaging agent etoposide or transfected with the indicated constructs before staining with DAPI (blue) and antibodies to 53BP1 (red). The fraction of cells containing more than three foci is indicated beneath each panel. The total number of cells analyzed were: etoposide, 81; GFP, 277; L^{wt}/R^{wt}, 214; L^{-}/R^{+}, 225. Scale bar, 10 μm. (**b**) ZFN expression as monitored by anti-FLAG western blot. The ZFNs migrate as a doublet ('FLAG' labeled bands) due to the presence of a 20-residue C-terminal tag on the L^{wt} and L^{-} ZFNs. TFIIB was monitored as a loading control.

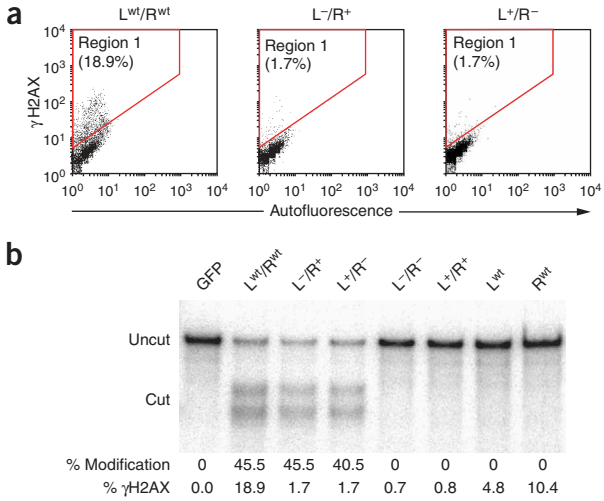

Figure 6 Heterodimer variants yield reduced DNA damage levels while retaining full activity for targeted gene modification. (**a**) Flow cytometry data for human K562 cells treated with the indicated constructs and stained with antibodies against γH2AX. The gate used to isolate γH2AX positive cells is shown in red. The percentage of cells falling within the gate for each sample is indicated. (**b**) DNA damage and gene modification rates in K562 cells transfected with the indicated constructs. Three days after transfection, each sample was monitored for DNA damage as described in **Figure 6a**. The observed percentage of γH2AX-positive cells is indicated at bottom. At the same time, gene-modification rates were monitored by PCR amplification of IL2Rγ exon 5 from the treated cell pool followed by digestion with the Surveyor nuclease (Transgenomic) and PAGE. This assay reveals gene modification by the appearance of digestion products ('Cut' bands), which may be quantified to yield the % gene modification indicated at the base of each lane. This assay has a limit of detection of ~1%, so the lack of signal observed in the L^{-}/R^{-} and L^{+}/R^{+} lanes suggests that there is at least a 40-fold preference for heterodimer function over homodimer function.

immunofluorescence[40–42]. Antibody-based detection of 53BP1 revealed that substitution of the heterodimer variants for the wild-type *Fok*I cleavage domain yielded a visually marked reduction in the number of 53BP1-stained foci in K562 cells to levels similar to those of background. This observation was confirmed by quantifying the fraction of cells containing multiple foci (**Fig. 5a**). This result was not due to poor expression by the L^{-}/R^{+} dimer, as western blot analysis confirmed comparable levels of nuclease expression for both the L^{wt}/R^{wt} and L^{-}/R^{+} combinations (**Fig. 5b**).

To confirm the 53BP1 results, we repeated these studies using antibodies for an alternative damage marker, phosphorylated histone H2AX (γH2AX), which is generated in response to DNA damage and forms foci at double-strand breaks[43–45]. Quantification relied on counting the fraction of antibody-stained cells by flow cytometry, with gates adjusted empirically to isolate the γH2AX signal from autofluorescence (**Fig. 6a**). We observed that substitution of the +/– heterodimers for the wild-type *Fok*I cleavage domain reduced the fraction of γH2AX-positive cells from 18.9% to 1.7% (**Fig. 6a**), indicating a substantial reduction in the number of off-target cleavage events induced by the variant domains. Moreover, when delivered individually both L^{wt} and R^{wt} yielded substantial numbers of γH2AX-positive cells indicating that off-target cleavage by homodimerization can occur with both ZFNs in this system (**Fig. 6b**).

Finally, to ensure that the improved genome-wide specificity was not achieved at the cost of efficacy at the intended target, we also monitored the IL2Rγ locus for ZFN-induced modification. In the absence of donor, gene repair through nonhomologous end-joining yields minor insertions and deletions at the site of a double-strand break[46]. Measuring the frequency of these events in a ZFN-treated cell pool using the surveyor nuclease assay provides a convenient means for gauging ZFN activity *in vivo*. The results of this assay (**Fig. 6b**) revealed comparable levels of gene modification in cells treated with L^{wt}/R^{wt}, L^{+}/R^{-} and L^{-}/R^{+}, indicating similar cellular activities of these ZFNs at the target locus. These results confirm that the heterodimer variants retain full activity for modification of the target locus while exhibiting considerably reduced levels of off-target cleavage.

DISCUSSION

In this study we have developed two complementary *Fok*I cleavage domain variants that together function as an obligate heterodimer and improve ZFN specificity *in vivo*. These variants should provide broad utility for improving ZFN specificity in gene-modification applications and have, in fact, already been used to improve the *in vivo* specificity of other ZFNs (K. Beumer and D. Carroll, University of Utah School of Medicine, personal communication). We anticipate at least three types of situations in which use of these variants will be advantageous. First, for therapeutic applications, the variants will offer an additional mechanism for minimizing off-target cleavage. Second, even for less safety-intensive applications, such as crop engineering, cell-line engineering and the construction of disease models, substitution of these variants for the wild-type *Fok*I cleavage domain will provide a straightforward means for improving ZFN properties (e.g., as an alternative to extensive optimization of the protein-DNA interface). Finally, we expect that our variants will prove useful in gene modification protocols requiring simultaneous cleavage at multiple targets (for example, to alter two specific loci at once). Cleavage at two targets would require cellular expression of four ZFNs, which would yield ten different active ZFN combinations. For such applications, substitution of our variants for the wild-type nuclease domain would eliminate the activity of six of these combinations and should dramatically reduce chances of off-target cleavage.

To our knowledge, reengineering of an endonuclease homodimer for obligate heterodimerization (and thus for improved functional specificity mediated by changes in the nuclease domain) has not been described previously. The general utility of this approach for improving cleavage specificity should be of interest to anyone wishing to engineer restriction enzymes for targeting chosen sequences[47–51]. The strategy that we have used in this study, consisting of the iterative design and screening of interface variants for progressively more heterodimeric behavior, should be generally applicable for converting other nucleases into obligate heterodimers. Our results suggest that this strategy may be quite robust: although our test panels were modest in size and chosen using relatively simple structure-based methods (**Supplementary Methods**), we nonetheless identified variants with the desired properties in just four design iterations. The use of more sophisticated design algorithms or the development of methods for testing larger numbers of variants would no doubt further improve both the outcomes and efficiency of this approach.

One aspect of this study that was initially unexpected was the magnitude of the decrease in global off-target cleavage events observed with our *Fok*I variants (**Figs. 5a** and **6b**). All else being equal, simply inactivating the homodimeric species might be expected to have a more modest effect. The most plausible explanation for this behavior is that in addition to inactivating the homodimeric species, our '+' and '–' variants interact with each other slightly more weakly than do two copies of the wild-type domain. To cleave DNA at the same

efficiency, a complex with a weaker affinity between the cleavage domains would require stronger interactions between each monomer and its target site. Thus, weakening the interaction between the two cleavage domains would have a disproportionate effect on low affinity off-target sites compared to the intended target. Consistent with this possibility, in preliminary *in vitro* cleavage studies, we observed that both the L^-/R^+ and L^+/R^- ZFN combinations cut DNA with apparent initial rates that were about threefold lower than a similar concentration of the L^{wt}/R^{wt} nucleases (data not shown). Thus it appears that our cell-based screen has yielded cleavage domain variants that not only heterodimerize, but also associate with a somewhat reduced affinity relative to the parental cleavage domain, yielding enhanced specificity *in vivo*.

Finally, we note that, to our knowledge, designed ZFNs have not previously been directly tested *in vivo* for off-target cleavage. Using two different approaches, we quantified DNA damage levels in cells treated with highly active ZFNs, and observed near-background amounts of genome-wide cleavage (**Figs. 5a** and **6b**). These studies, in combination with recent publications, reemphasize the potential of ZFNs as efficient and specific gene-modification agents. Moreover, as an initial genome-wide test of any designed sequence-specific nuclease, these studies provide a benchmark by which to gauge the performance of other such agents. Several alternative design platforms have been used to construct nucleases with targeted specificities, and it will be interesting to see how they perform in cellular DNA damage assays. More generally, our results indicate that properly designed ZFNs can cleave DNA *in vivo* with high selectivity, supporting their potential in diverse applications including gene-modification therapy.

METHODS

Molecular modeling. Modeling of the FokI dimer interface started with 2FOK.pdb[38] and was performed with VMD[52] and Swiss-Pdb Viewer[53] (**Supplementary Methods**). Models of the zinc-finger domains were built from fingers two and three of the Zif268 co-crystal structure (1AAY.pdb)[10], whereas the *Fok*I cleavage domain dimer is from 2FOK.pdb[38], and was docked to DNA based upon the protein-DNA interaction of *Bam*HI (2BAM.pdb)[54].

Zinc-finger nuclease constructs. The amino acid sequences of the L^{wt} and R^{wt} ZFNs are identical to those described as ZFP-L* and ZFP-R* (ref. 5). Mutations were introduced into the region encoding the *Fok*I cleavage domain of each ZFN using the QuikChange mutagenesis kit (Invitrogen). The L^{wt} and R^{wt} constructs contain residues 384 to 579 of the *Fok*I endonuclease[38]. The complete amino acid sequences of the L^{wt} and R^{wt} nucleases and the *Fok*I domain variants described in **Figure 2d** are given in **Supplementary Figure 2** online.

Reporter gene correction in HEK293 cell line. A defective eGFP reporter gene containing a frameshift mutation and a fragment of the ZFN-targeted stretch of the IL2Rγ gene was generated as described[3], cloned into the pcDNA4/TO vector (Invitrogen) and stably introduced into HEK293 T-Rex cells (Invitrogen). A clonal line bearing a single integrated copy of this gene was identified and used for all experiments. The donor plasmid for correcting the defective eGFP gene was generated by cloning the functional eGFP gene (BD BioSciences) into pcDNA4/TO. To prevent GFP expression from the donor construct, the first 12 bp and start codon were removed. A 1.5-kb fragment was then generated by PCR using the following primers: 5′-GGCGAGGAGCTGTTCAC-3′, 5′-TGCATACTTCTGCCTGC-3′ (the reverse primer anneals to the pcDNA4/TO vector backbone) and cloned into pCR4Topo (Invitrogen). The reporter line was grown according to the manufacturer's instructions for HEK293 T-Rex cells and transfected with the two ZFN plasmids and donor construct using Lipofectamine 2000 (Invitrogen). Gene correction frequency was measured 4 d after transfection with a Guava Easycyte single-cell analysis system. Fluorescence at both 525 nm and 583 nm was independently measured to separate the GFP signal from autofluorescence (the GFP signal was stronger in the 525-nm channel, whereas autofluorescence was similar in both channels). Note that, as with other similar reporter lines, the absolute gene modification frequency observed in this experiment is lower than that seen at endogenous loci due at least in part to the sequence divergence between the donor and target caused by the inserted exogenous DNA stretch.

Endogenous gene modification in K562 cells. The donor plasmid used to introduce a *Bsr*BI restriction site into exon 5 of the IL2Rγ gene has been previously described[5]. For gene editing, K562 cells (ATCC) were grown to confluence in RPMI media (Gibco) supplemented with 10% FBS and split 1:4 1 d before transfection. Approximately 2×10^6 cells were nucleofected with 5 μg of each ZFP plasmid and 25 μg of donor plasmid using Cell Line Nucleofector Kit V and Program T16 (Amaxa Biosystems) according to the manufacturer's protocol. Cells were collected 48 h after transfection. DNA extraction was performed with the DNeasy Tissue Kit (Qiagen) according to the supplier's protocol. Percentage of gene editing was determined using a PCR-based assay as previously described[5] except that the PCR step used a reduced amount of genomic DNA (8 ng), was run for 24-cycles and included 16 μCi each of alpha-^{32}P-dCTP and alpha-^{32}P-dATP.

***In vitro* DNA cleavage assay.** Two complementary synthetic 69-mer oligonucleotides were annealed and cloned into the topo-blunt vector (Invitrogen) to generate each of the three sequences indicated in **Figure 4a**. The ZFN target sites in each construct were verified by DNA sequencing. The verified templates were then amplified with the M13 forward and M13 reverse primers included with the topo blunt kit and using the recommended PCR protocol except that the reaction was spiked with 5 μCi each of alpha-^{32}P-dCTP and alpha-^{32}P-dATP. The unincorporated nucleotides were removed with a G50 spin column (GE healthcare). To give a DNA concentration in the final reaction of ~1 nM, the resulting mixture was diluted 100-fold in *Fok*I buffer consisting of 20 mM Tris-HCl pH 8.5, 150 mM NaCl, 2 mM $MgCl_2$, 5% (vol/vol) glycerol, 10 μM $ZnCl_2$, 0.5 mg/ml BSA and 1 mM DTT. The ZFNs were expressed using the TnT-Quick coupled transcription/translation system (Promega) according to the manufacturer's recommendations, except that the incubation time was increased to 2 h. The appropriate ZFNs were mixed and diluted with 1 volume of *Fok*I buffer. A mock reaction without any starting template and diluted with 1 volume of *Fok*I buffer was used to make a sixfold dilution of the L^{wt}/R^{wt} mixture to approximately normalize the activity of all three ZFN mixtures on the LR target. Equal volumes of diluted target and diluted protein mixture were mixed and incubated at 37 °C for 2 h. The radiolabeled DNA was then extracted with a phenol/chloroform mixture and analyzed by PAGE followed by quantification with a phosphorimager system (Molecular Dynamics).

Measurement of double-strand breaks in ZFN-treated cells. K562 cells were grown as specified above and transfected by Nucleofection (Amaxa Biosystems) according to the manufacturer's protocol (Solution V, Program T16). The ZFNs were delivered on separate plasmids except for the experiment shown in **Figure 5**; in this case both ZFNs were expressed from a single transcript separated by the coding region of the 2A peptide, which yields each ZFN as a separate protein product also containing a portion of the 2A peptide[55]. This results in the fusion of an extra 20-residue C-terminal tag to the L ZFN, which yields migration of the L/R ZFN dimers as doublets in the western blot in **Figure 5b**. For 53BP1 immunocytochemistry, cells were collected 24 h after nucleofection and used to prepare slides by cytospin (Thermo Scientific). Cells were fixed with cold methanol and treated with 0.5% Triton X-100 buffer (0.5% Triton X-100, 1% BSA, 0.02% NaN3, 98.5% PBS) at 22 °C for 5 min. Cells were incubated with 5% goat serum to block nonspecific staining. Cells were then incubated with anti-53BP1 rabbit polyclonal antibodies (Bethyl Laboratories) followed by incubation with Alexa Fluor 594-conjugated secondary antibodies (Invitrogen-Molecular Probes) in the presence of 2.5 μg/ml of DAPI (Sigma) to counterstain cell nuclei. Slides were mounted and examined under an immunofluorescence microscope and images were acquired with a digital camera connected to the microscope.

For γH2AX flow cytometric analysis, cells were collected 2 d after nucleofection, fixed with 2% paraformaldehyde-PBS and permeabilized with

perm/wash buffer (0.05% Saponin, 2.5% FBS, 0.02% NaN3, 97.5% PBS). Cells were then incubated with freshly labeled anti-γH2AX monoclonal antibody (Upstate) or freshly labeled control mouse IgG1 (Invitrogen) as a negative control; both antibodies were labeled with a Zenon R-phycoerythrin (PE) rabbit IgG-labeling kit (Invitrogen-Molecular Probes) according to the manufacturer's instructions. Stained cells were analyzed using a Guava Easycyte single cell analysis system. Fluorescence at both 525 nm and 583 nm was independently measured to separate the phycoerythrin signal from autofluorescence (compensation was used isolate the phycoerythrin signal in the 583-nm channel).

Surveyor nuclease assay for quantification of gene disruption. The Surveyor nuclease (Transgenomic) efficiently and specifically cleaves DNA duplexes at the sites of distortions created by either bulges or mismatches[56]. This property enables its use to detect and quantify mutation rates even when the sequences of the modified products are variable and/or unknown. We have adapted protocols using this enzyme for quantification of the minor deletions and insertions typically induced by nonhomologous end-joining–mediated break repair. Our procedure is as follows: (i) first, genomic DNA is isolated from the nuclease-treated cell pool (Qiagen DNeasy kit); (ii) next, the targeted region (in this case IL2Rγ exon 5) is PCR-amplified using the AccuPrime PCR kit from Invitrogen and 50–200 ng of genomic DNA as template. ^{32}P-labeled dNTPs (5 μCi dATP and 5 μCi dCTP) are added to enable radiographic visualization of the amplicon; (iii) the PCR reaction is cleaned up using a G-50 spin column and the purified amplicons diluted by a factor of up to 1:10 into a buffer of 10 mM Tris/100 mM NaCl pH 8 to equalize counts; (iv) a melting/annealing step is then performed using a temperature program of: 95 °C/10 min; 95 °C to 85 °C (–2 °C/s); 85 °C to 25 °C (–0.1 °C/s); 4 °C hold. This step randomly reassorts mutant and wild-type DNA strands with the consequent formation of distorted duplex DNA; (v) reannealed duplex (10 μl) is combined with 1 μl of Surveyor enzyme, 1 μl 10× AccuPrime PCR buffer II and water to a total volume of 20 μl; (vi) samples are incubated at 42 °C for 20 min and electrophoresed on a 10% TBE polyacrylamide gel (BioRad); (vi) the dried gel is exposed and quantified using a Phosphorimager system (Molecular Dynamics). At low gene-modification efficiencies ($< \sim 10\%$) the reannealing procedure (step iv) converts each mutated amplicon into two distorted duplexes through reassociation with wild-type strands, which are in excess. In this range one may therefore estimate gene-modification rates by using the approximation: fractional modification = fraction of cleaved bands/2. At higher levels of modification, such as those in **Figure 6**, a more general and exact equation is required to correct for the fact that mutated strands do not exclusively anneal with wild-type sequence (a nontrivial amount of mutated:mutated duplexes will form). The relevant equation is: fractional modification = $(1- (1- (\text{fraction cleaved}))^{1/2})$. In control studies comparing gene modification rates deduced by the Surveyor assay and direct sequencing the two methods have generally agreed to within a factor of 0.05.

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

ACKNOWLEDGMENTS

We thank Judy Campisi for support and helpful discussions, Yann Jouvenot, Sheldon Augustus, Danny Xia and Lei Zhang for assistance with protocols and constructs, Dana Carroll for comments on the manuscript, Fyodor Urnov for helpful discussions and Edward Lanphier for encouragement and support. This research was supported in part by grant no. 7ONANB4H3006 from the Advanced Technology Program (US Department of Energy).

AUTHOR CONTRIBUTIONS

J.C.M and C.O.P. conceived the project; J.C.M., M.C.H., P.D.G., C.O.P. and E.J.R. designed experiments; J.C.M., M.C.H., J.W., D.Y.G., Y.-L.L., I.R., C.M.B., A.J.W., N.S.W and K.A.K. performed the experiments; J.C.M., M.C.H., P.D.G., C.O.P. and E.J.R. wrote the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturebiotechnology/.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Kim, Y.G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **93**, 1156–1160 (1996).
2. Smith, J., Berg, J.M. & Chandrasegaran, S. A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucleic Acids Res.* **27**, 674–681 (1999).
3. Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
4. Bibikova, M., Beumer, K., Trautman, J.K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
5. Urnov, F.D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
6. Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K. & Carroll, D. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics* **172**, 2391–2403 (2006).
7. Moehle, E.A. *et al.* Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl. Acad. Sci. USA* **104**, 3055–3060 (2007).
8. Liang, F., Han, M., Romanienko, P.J. & Jasin, M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc. Natl. Acad. Sci. USA* **95**, 5172–5177 (1998).
9. Elrod-Erickson, M., Benson, T.E. & Pabo, C.O. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger-DNA recognition. *Structure* **6**, 451–464 (1998).
10. Elrod-Erickson, M., Rould, M.A., Nekludova, L. & Pabo, C.O. Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. *Structure* **4**, 1171–1180 (1996).
11. Miller, J.C. & Pabo, C.O. Rearrangement of side-chains in a Zif268 mutant highlights the complexities of zinc finger-DNA recognition. *J. Mol. Biol.* **313**, 309–315 (2001).
12. Greisman, H.A. & Pabo, C.O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275**, 657–661 (1997).
13. Segal, D.J., Dreier, B., Beerli, R.R. & Barbas, C.F., III. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl. Acad. Sci. USA* **96**, 2758–2763 (1999).
14. Hurt, J.A., Thibodeau, S.A., Hirsh, A.S., Pabo, C.O. & Joung, J.K. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc. Natl. Acad. Sci. USA* **100**, 12271–12276 (2003).
15. Pabo, C.O., Peisach, E. & Grant, R.A. Design and selection of novel Cys2His2 zinc finger proteins. *Annu. Rev. Biochem.* **70**, 313–340 (2001).
16. Isalan, M., Klug, A. & Choo, Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat. Biotechnol.* **19**, 656–660 (2001).
17. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374 (2001).
18. Lloyd, A., Plaisier, C.L., Carroll, D. & Drews, G.N. Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **102**, 2232–2237 (2005).
19. Wright, D.A. *et al.* High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* **44**, 693–705 (2005).
20. Bibikova, M., Golic, M., Golic, K.G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
21. Morton, J., Davis, M.W., Jorgensen, E.M. & Carroll, D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proc. Natl. Acad. Sci. USA* **103**, 16370–16375 (2006).
22. Bitinaite, J., Wah, D.A., Aggarwal, A.K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. USA* **95**, 10570–10575 (1998).
23. Vanamee, E.S., Santagata, S. & Aggarwal, A.K. FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.* **309**, 69–78 (2001).
24. Smith, J. *et al.* Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**, 3361–3369 (2000).
25. Mani, M., Smith, J., Kandavelou, K., Berg, J.M. & Chandrasegaran, S. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochem. Biophys. Res. Commun.* **334**, 1191–1197 (2005).
26. Kaczorowski, T., Skowron, P. & Podhajska, A.J. Purification and characterization of the FokI restriction endonuclease. *Gene* **80**, 209–216 (1989).
27. Richardson, C., Elliott, B. & Jasin, M. Chromosomal double-strand breaks introduced in mammalian cells by expression of I-Sce I endonuclease. *Methods Mol. Biol.* **113**, 453–463 (1999).
28. Tan, S. *et al.* Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity. *Proc. Natl. Acad. Sci. USA* **100**, 11997–12002 (2003).
29. Bartsevich, V.V., Miller, J.C., Case, C.C. & Pabo, C.O. Engineered zinc finger proteins for controlling stem cell fate. *Stem Cells* **21**, 632–637 (2003).
30. Beerli, R.R., Dreier, B. & Barbas, C.F. 3rd Positive and negative regulation of endogenous genes by designed transcription factors. *Proc. Natl. Acad. Sci. USA* **97**, 1495–1500 (2000).

31. Guan, X. *et al.* Heritable endogenous gene regulation in plants with designed polydactyl zinc finger transcription factors. *Proc. Natl. Acad. Sci. USA* **99**, 13296–13301 (2002).
32. O'Shea, E.K., Lumb, K.J. & Kim, P.S. Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658–667 (1993).
33. Zhu, Z., Presta, L.G., Zapata, G. & Carter, P. Remodeling domain interfaces to enhance heterodimer formation. *Protein Sci.* **6**, 781–788 (1997).
34. Atwell, S., Ridgway, J.B., Wells, J.A. & Carter, P. Stable heterodimers from remodeling the domain interface of a homodimer using a phage display library. *J. Mol. Biol.* **270**, 26–35 (1997).
35. Nohaile, M.J., Hendsch, Z.S., Tidor, B. & Sauer, R.T. Altering dimerization specificity by changes in surface electrostatics. *Proc. Natl. Acad. Sci. USA* **98**, 3109–3114 (2001).
36. Havranek, J.J. & Harbury, P.B. Automated design of specificity in molecular recognition. *Nat. Struct. Biol.* **10**, 45–52 (2003).
37. Bolon, D.N., Grant, R.A., Baker, T.A. & Sauer, R.T. Specificity versus stability in computational protein design. *Proc. Natl. Acad. Sci. USA* **102**, 12724–12729 (2005).
38. Wah, D.A., Bitinaite, J., Schildkraut, I. & Aggarwal, A.K. Structure of FokI has implications for DNA cleavage. *Proc. Natl. Acad. Sci. USA* **95**, 10564–10569 (1998).
39. Joachimiak, L.A., Kortemme, T., Stoddard, B.L. & Baker, D. Computational design of a new hydrogen bond network and at least a 300-fold specificity switch at a protein-protein interface. *J. Mol. Biol.* **361**, 195–208 (2006).
40. Schultz, L.B., Chehab, N.H., Malikzay, A. & Halazonetis, T.D. p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks. *J. Cell Biol.* **151**, 1381–1390 (2000).
41. Anderson, L., Henderson, C. & Adachi, Y. Phosphorylation and rapid relocalization of 53BP1 to nuclear foci upon DNA damage. *Mol. Cell. Biol.* **21**, 1719–1729 (2001).
42. Rappold, I., Iwabuchi, K., Date, T. & Chen, J. Tumor suppressor p53 binding protein 1 (53BP1) is involved in DNA damage-signaling pathways. *J. Cell Biol.* **153**, 613–620 (2001).
43. Rogakou, E.P., Pilch, D.R., Orr, A.H., Ivanova, V.S. & Bonner, W.M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
44. Rogakou, E.P., Boon, C., Redon, C. & Bonner, W.M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol.* **146**, 905–916 (1999).
45. Stiff, T. *et al.* ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer Res.* **64**, 2390–2396 (2004).
46. Jeggo, P.A. DNA breakage and repair. *Adv. Genet.* **38**, 185–218 (1998).
47. Ashworth, J. *et al.* Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441**, 656–659 (2006).
48. Sussman, D. *et al.* Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J. Mol. Biol.* **342**, 31–41 (2004).
49. Arnould, S. *et al.* Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J. Mol. Biol.* **355**, 443–458 (2006).
50. Chames, P. *et al.* In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination. *Nucleic Acids Res.* **33**, e178 (2005).
51. Epinat, J.C. *et al.* A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* **31**, 2952–2962 (2003).
52. Humphrey, W., Dalke, A. & Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
53. Guex, N. & Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
54. Viadiu, H. & Aggarwal, A.K. The role of metals in catalysis by the restriction endonuclease BamHI. *Nat. Struct. Biol.* **5**, 910–916 (1998).
55. Szymczak, A.L. *et al.* Correction of multi-gene deficiency *in vivo* using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.* **22**, 589–594 (2004).
56. Qiu, P. *et al.* Mutation detection using Surveyor nuclease. *Biotechniques* **36**, 702–707 (2004).
57. Kortemme, T. *et al.* Computational redesign of protein-protein interaction specificity. *Nat. Struct. Mol. Biol.* **11**, 371–379 (2004).