Published online 7 July 2009

Nucleic Acids Research, 2009, Vol. 37, No. 16 ***5405–5419***
doi:10.1093/nar/gkp548

Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease

Sylvestre Grizot[1], Julianne Smith[2], Fayza Daboussi[1], Jesús Prieto[3], Pilar Redondo[3], Nekane Merino[4], Maider Villate[4], Séverine Thomas[1], Laetitia Lemaire[2], Guillermo Montoya[3], Francisco J. Blanco[4], Frédéric Pâques[1,2,*] and Philippe Duchateau[1]

[1]Cellectis S.A., [2]Cellectis Genome Surgery, 102 Avenue Gustave Roussel, 93235 Romainville, France, [3]Structural Biology and Biocomputing Program, Spanish National Cancer Center (CNIO), Macromolecular Crystallography Group, Melchor Fdez. Almagro 3, 28029 Madrid and [4]Structural Biology Unit, CIC bioGUNE, Parque Tecnológico de Vizcaya, Edificio 800 48160, Derio, Spain

Received May 5, 2009; Revised June 10, 2009; Accepted June 11, 2009

ABSTRACT

Sequence-specific endonucleases recognizing long target sequences are emerging as powerful tools for genome engineering. These endonucleases could be used to correct deleterious mutations or to inactivate viruses, in a new approach to molecular medicine. However, such applications are highly demanding in terms of safety. Mutations in the human *RAG1* gene cause severe combined immunodeficiency (SCID). Using the I-CreI dimeric LAGLIDADG meganuclease as a scaffold, we describe here the engineering of a series of endonucleases cleaving the human *RAG1* gene, including obligate heterodimers and single-chain molecules. We show that a novel single-chain design, in which two different monomers are linked to form a single molecule, can induce high levels of recombination while safeguarding more effectively against potential genotoxicity. We provide here the first demonstration that an engineered meganuclease can induce targeted recombination at an endogenous locus in up to 6% of transfected human cells. These properties rank this new generation of endonucleases among the best molecular scissors available for genome surgery strategies, potentially avoiding the deleterious effects of previous gene therapy approaches.

INTRODUCTION

Meganucleases are endonucleases that recognize large (>12 bp) DNA sequences. They are often referred to as homing endonucleases (HE) because the proteins are encoded by mobile introns or inteins which propagate via a process called homing. Their target sequence is usually found in homologous alleles lacking the intron or intein. Cleavage of the target DNA sequence stimulates a homologous recombination event resulting in the insertion of the mobile genetic element into the DNA cleavage site (1). The homing endonuclease I-SceI has also been shown to stimulate targeted recombination in immortalized mammalian cells (2,3) and in various cell types and organisms (4). These results, together with the high specificity of meganucleases, make these enzymes a tool of choice for genome engineering. Until recently, the major limitation of this approach was the absence of any cleavage sequences for natural meganucleases in the vicinity of the locus of interest, and three different strategies have been envisioned for the generation of engineered artificial endonucleases cleaving chosen sequences (5). Chimeric endonucleases resulting from the fusion of engineered zinc-fingers domains with the catalytic domain of the *FokI* restriction enzyme have been used to induce targeted recombination in endogenous genes in human cells, Drosophila, plants, fish and other organisms (6–12). In a totally different approach, triplex forming oligonucleotides (TFOs) have been conjugated with chemical DNA cleavers (13) or restriction enzymes (14) to target their cleavage activity to specific long sequences. However, the recent development of adapted protein-engineering strategies (15–23) has opened the way to the development of tailored endonucleases derived from natural HE.

HEs are classified in five families, of which the LAGLIDADG family (named after a conserved peptide motif) is the largest and best characterized one (24). Proteins with one LAGLIDADG motif in their primary sequence, such as I-CreI or I-MsoI form homodimers and cleave palindromic or pseudo-palindromic sequences,

*To whom correspondence should be addressed. Tel: 33 1 41 83 99 14; Fax: 33 1 41 83 99 03; Email: paques@cellectis.com

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

whereas HE containing two such motifs, such as I-SceI or I-DmoI, are single-chain proteins, that cleave non-palindromic targets (25). The structures of several members of this family have been solved (26–30), showing a strong conservation of the core structure. The LAGLIDADG motifs are part of the conserved α-helices constituting the main interface between the two monomers or the two domains of a single-chain protein. Each of the monomers or the domains adopts a characteristic αββαββα fold and DNA recognition and binding is dependent on the four-stranded antiparallel β-sheets, which form a saddle on the major groove of the DNA helix. In addition, structures of the complexes formed between the endonuclease and cleaved or uncleaved forms of the DNA target have provided insight into the catalytic mechanism underlying hydrolysis of the two phosphodiester bonds (31).

We have already described how, by adopting a semi-rational combinatorial strategy, we were able to design I-CreI variant proteins that were able to cleave DNA targets of interest, including DNA sequences located in the human *RAG1* and *XPC* genes in particular (15,32,33). The non-palindromic DNA sequence is cleaved following the co-expression of two I-CreI variants, each of them recognizing one-half of the DNA target. The two proteins can associate as a heterodimer, which binds and cleaves the DNA target. Nevertheless, the co-production of two I-CreI derived proteins results in the presence of three molecular species in the cell: the heterodimer and both homodimers. The two homodimers may be considered to be by-products with the potential to affect the overall meganuclease specificity, a particularly crucial aspect for therapeutic applications (4). Two possible strategies for overcoming the problem of homodimer formation could be envisaged. The first involves engineering the protein interface between the two monomers to impair the formation of functional homodimers and favor heterodimer formation. This strategy for designing an obligate heterodimer has already been described for zinc-finger nucleases (34,35) and for I-CreI derived meganucleases (36). This strategy is still limited by the need to co-transfect cells with two plasmids, making it impossible to control the stoichiometry of the two monomers. The second strategy would involve mimicking natural evolution by creating a single-chain molecule. This approach has already been successfully applied to I-CreI and I-MsoI proteins (37,38). The single-chain design would not only abolish homodimer activity, but would also facilitate transfection and resolve stoichiometry problems. We describe here the engineering of obligate heterodimers and single-chain variants of I-CreI for targeting the human *RAG1* gene. These two designs result in meganucleases with high levels of activity *in vivo* against the target sequence and an absence of homodimer activity.

The lymphoid-specific Rag1 protein is essential for the V(D)J recombination process, which generates the diversity of the immune repertoire. The disruption of Rag1 function leads to a severe combined immunodeficiency (SCID), with a complete absence of mature B and T cells (39). The induction of a homologous recombination mechanism following the creation of a DNA double strand break (DSB) by a specific meganuclease could correct *RAG1* mutations and lead to reversion to a normal phenotype. However, this approach to genome engineering for therapeutic purposes requires the highest possible level of meganuclease specificity. Therefore, we designed obligate heterodimer and single-chain versions of the RAG1 meganuclease. The protein activity and toxicity of each RAG meganuclease were assessed in various assays, and both designs appeared to result in significantly lower levels of toxicity, while preserving or even enhancing activity. The best properties were displayed by single-chain molecules, which were used to induce high levels of recombination (up to 6%) in the endogenous human *RAG1* gene in 293 cells, demonstrating the full potential of this technology. Thus the single chain appears to be the scaffold of choice to undertake the genomic correction of mutations present in the human genome for gene therapy in SCID and other monogenic diseases.

MATERIALS AND METHODS

Construction of target clones

The 22-bp LR target found in the human *RAG1* gene (TGTTCTCAGGTACCTCAGCCAG), the two palindromic derivative targets LL (TGTTCTCAGGTACCTGAG AACA) and RR (CTGGCTGAGGTACCTCAGCCAG) were cloned as follows: oligonucleotides containing the target site (Proligo) were amplified by PCR to generate double-stranded target DNA and then inserted into reporter vectors with the Gateway protocol (Invitrogen): the yeast vector pFL39-ADH-LACURAZ and the mammalian vector pcDNA3.1-LAACZ, both previously described and containing an I-SceI target site as a control. Yeast reporter vectors were used to transform *Saccharomyces cerevisiae* strain FYBL2-7B (MAT a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202).

Construction of the single-chain molecule

One single-chain molecule contains two almost identical sequences, encoding the I-CreI derived V2 and V3 mutants. The two coding sequences are different, displaying only 72% nucleic acid sequence identity, thereby avoiding problems relating to construct stability. The N-terminal half of the single-chain molecules contains a full length I-CreI molecule of 163 residues, whereas the C-terminal sequence lacks the first five residues. Various linkers were used to connect the two mutant proteins (Table 1).

Meganuclease production

The ORFs encoding the various meganucleases were amplified by PCR and inserted into the 2 μm-based replicative vector pCLS542, harboring the *LEU2* gene. *Saccharomyces cerevisiae* strain FYC2-6A (MATα, trp1Δ63, leu2Δ1, his3Δ200) was then transformed with the vector using a high-efficiency lithium acetate transformation protocol. The various meganuclease genes were also inserted into the mammalian vector pcDNA3.1. The genes were expressed under the control of a CMV promoter. Protein production in CHO-K1 cells was

monitored by adding an HA (YPYDVPDYA) or a S-Tag (KETAAAKFERQHMDS) epitope to the C-terminus of the protein. Cells were harvested 48 h after the transfection of the cells using the Polyfect reagent (Qiagen) and directly solubilized in Laemmli buffer (100 µl of buffer for 10^6 cells). The equivalent of 10^5 cells was loaded on a SDS–PAGE gel and probed by western blot using an anti-HA antibody (Roche) or an anti-S-Tag antibody (Santa Cruz).

Purification of meganucleases

For the production of heterodimeric I-CreI derivatives, the ORFs of each of the monomers were inserted into the CDFDuet-1 vector (Novagen) with a 6xHis tag or a Strep tag at the C-terminus. The double-tagged heterodimers were produced and purified as previously described (30). The protein was concentrated to about 15 mg/ml, flash-frozen in liquid nitrogen and stored at −80°C. Protein concentration was determined by measuring absorbance at 280 nm. Sample purity was checked by SDS–PAGE and heterodimer formation was assessed by western blotting with an anti-His or an anti-Strep tag antibody and confirmed by mass spectrometry. All the purified proteins were found to have adopted a structure similar to that of the wild type (as shown by circular dichroism and NMR) and to be dimeric in solution (as shown by analytical ultracentrifugation: data not shown).

The sequence encoding the single-chain scV3-V2(G19S) protein with a 6xHis tag at the C-terminus was inserted into the pET24d(+) vector and expressed in *Escherichia coli* Rosetta(DE3)pLysS cells (Novagen) grown in LB supplemented with kanamycin and chloramphenicol. Induction with IPTG for 5 h at 37°C or for 15 h at 20°C yielded high levels of expression, however, after sonication in lysis buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 5% glycerol and protease inhibitors (Complete EDTA-Free tablets, Roche) and ultracentrifugation at $20\,000 \times g$ for 1 h, the protein was found exclusively in the insoluble fraction as determined by western blots with an anti-His antibody. The protein was therefore purified under denaturing conditions by first solubilizing it in lysis buffer supplemented with 8 M urea. After clarification by ultracentrifugation (2 h at $40\,000 \times g$) the sample was applied onto a column packed with Q-Sepharose XL resin (GE Healthcare) equilibrated with the same buffer. This purification step separated all the nucleic acids (retained in the column) from the protein and improved the performance of subsequent purification steps. The protein was recovered from the flow through on a Co^{2+}-loaded HiTrap Chelating HP 5 ml column (GE Healthcare) equilibrated with the lysis buffer plus 8 M urea. After sample loading and column washing the protein was eluted with the same buffer supplemented with 0.5 M imidazol. Protein-rich fractions (as determined by SDS–PAGE) were collected and allowed to undergo refolding, through a 20-fold dilution (drop by drop) with 20 mM sodium phosphate pH 6.0, 300 mM NaCl at 4°C (final protein concentration of 0.13 mg/ml). The refolded protein was loaded onto a 5 ml HiTrap heparin column equilibrated with the same buffer and eluted with a gradient to 1 M NaCl. The fractions containing the pure protein were pooled, concentrated up to 1.4 mg/ml (35.6 µM, determined by absorbance at 280 nm) and were either used immediately or flash frozen in liquid nitrogen and stored at −80°C. The identity of the protein was confirmed by mass spectrometry which showed that the initial methionine was absent from the purified polypeptide chain. The purified protein was found to be folded into a structure similar to that of the wild type by circular dichroism and NMR (data not shown), and to be monomeric in solution by analytical gel filtration.

In vitro cleavage assay conditions

Cleavage assays were performed at 37°C in 10 mM Tris–HCl pH 8, 50 mM NaCl, 10 mM $MgCl_2$. The target concentration was 2 nM (target substrates in plasmid pGEM-T linearized with XmnI) and the protein concentrations were 120 (for cleavage detection) and 120, 90, 60, 40, 30, 20, 10, 7.5, 5, 3.5, 2, 1, 0.5, 0.25 and 0 nM for the quantification of cleavage efficiency, in a final reaction volume of 25–50 µl. Reactions were stopped after 1 h, adding 5 µl of 45% glycerol, 95 mM EDTA (pH 8), 1.5% (w/v) SDS, 1.5 mg/ml proteinase K and 0.048% (w/v) bromophenol blue (6× stop buffer) and incubating at 37°C for 30 min. The products of the reaction were subjected to electrophoresis in a 1% agarose gel. The gels were stained with SYBR Safe DNA gel staining kits (Invitrogen) and the intensity of the bands observed upon illumination with UV light was determined with the ImageJ software (http://rsb.info.nih.gov/ij/). The linearized target plasmid was 3 kb in size and yielded two smaller bands, of 2 kb and 1 kb, upon cleavage with the meganuclease. The percentage of cleavage was calculated with the following equation: $\%\ \text{cleavage} = 100 \times (I_{2\,\text{kb}} + I_{1\,\text{kb}})/(I_{3\,\text{kb}} + I_{2\,\text{kb}} + I_{1\,\text{kb}})$, where $I_{1\,\text{kb}}$, $I_{2\,\text{kb}}$ and $I_{3\,\text{kb}}$ are the intensities of the 1, 2 and 3 kb bands, respectively.

Mating of meganuclease expressing clones and screening in yeast

A colony gridder (QpixII, Genetix) was used for the mating of yeast strains. Mutants were gridded on nylon filters covering YPD plates, using a high gridding density (about 20 spots/cm^2). A second gridding process was performed on the same filters for the spotting of a second layer consisting of reporter-harboring yeast strains for each target. Membranes were placed on solid agar YPD rich medium, and incubated overnight at 30°C, to allow mating. The filters were then transferred onto synthetic medium, lacking leucine and tryptophan, with galactose (2%) as the carbon source (and with G418 for co-expression experiments), and incubated for 5 days at 37°C, to select for diploids carrying the expression and target vectors. Filters were then placed on solid agarose medium with 0.02% X-Gal in 0.5 M sodium phosphate buffer, pH 7.0, 0.1% SDS, 6% dimethyl formamide (DMF), 7 mM β-mercaptoethanol, 1% agarose, and incubated at 37°C, to monitor β-galactosidase activity. Filters were scanned and quantified with proprietary software. Beta-galactosidase activity is directly associated with the efficiency of homologous recombination. Experiments

using several purified I-CreI mutants with various recombination activities in yeast have shown that the recombination efficiency quantified in yeast is directly correlated with the cleavage activity *in vitro* (unpublished data).

Extrachromosomal assay in CHO-K1 cells

CHO-K1 cells were transfected with the meganuclease expression vectors and the reporter plasmid, in the presence of Polyfect transfection reagent in accordance with the manufacturer's protocol (Qiagen) The culture medium was removed 72 h after transfection, and 150 µl of lysis/detection buffer was added for β-galactosidase liquid assay (typically, for 1 l of buffer, we used 100 ml of lysis buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.1% Triton X100, 0.1 mg/ml BSA, protease inhibitors), 10 ml of Mg 100× buffer ($MgCl_2$ 100 mM, 2-mercaptoethanol 35%), 110 ml of an 8 mg/ml solution of ONPG and 780 ml of 0.1 M sodium phosphate pH 7.5). After incubation at 37°C, we measured optical density at 420 nm. The entire process was performed on an automated Velocity11 BioCel platform.

Immuno-precipitation experiments

CHO-K1 cells were transfected with the meganuclease expression vectors in the presence of Polyfect transfection reagent according to the manufacturer's (Qiagen) protocol. Cells were harvested 48 h after transfection. We avoided the disruption of heterodimers due to the use of stringent buffers such RIPA, by resuspending cells in PBS supplemented with a mixture of protease inhibitors (Santa Cruz) and lysing them by successive freeze-thaw cycles. After centrifugation (10 min at $13\,000 \times g$), the supernatant was recovered and the total protein concentration was determined by the BCA assay. Volumes were adjusted such that the protein concentration was identical (0.7 mg/ml) for the V2/V3 heterodimer and the $V2^+/V3^-$ obligate heterodimer. Both supernatants were then incubated twice, for 12 h each, with 15 µl (equivalent to 7.5 µg of antibody) of an anti-HA antibody coupled to agarose at 4°C. The sample was then centrifuged (2 min at $8000 \times g$) and 10 µl of the supernatant were subjected to SDS–PAGE and western blotting.

Targeting of a chromosomal reporter gene in CHO-K1 cells

CHO-K1 cell lines harbouring the reporter system were seeded at a density of 2×10^5 cells per 10 cm dish in complete medium [Kaighn's modified F-12 medium (F12-K), supplemented with 2 mM L-glutamine, penicillin (100 IU/ml), streptomycin (100 µg/ml), amphotericin B (Fongizone) (0.25 µg/ml) (Invitrogen-Life Science) and 10% FBS (Sigma-Aldrich Chimie)]. The next day, cells were transfected in the presence of Polyfect transfection reagent (Qiagen). Cells were co-transfected with 2 µg of LacZ repair matrix vector and various amounts of meganuclease expression vector. After 72 h of incubation at 37°C, cells were fixed by incubation in 0.5% glutaraldehyde at 4°C for 10 min, washed twice in 100 mM phosphate buffer supplemented with 0.02% NP40 and stained with the following staining buffer [10 mM phosphate buffer, 1 mM $MgCl_2$, 33 mM potassium hexacyanoferrate (III), 33 mM potassium hexacyanoferrate (II), 0.1% (w/v) X-Gal]. Plates were incubated overnight at 37°C and examined under a light microscope to determine the number of LacZ-positive cells. The frequency of LacZ repair is expressed as a percentage and is calculated as the number of LacZ-positive foci divided by the number of transfected cells (5×10^5), corrected for the transfection efficiency.

Cell survival assay

The CHO-K1 cell line was used to seed plates at a density of 2×10^5 cells per 10 cm dish. The next day, various amounts of meganuclease expression vectors and a constant amount of GFP-encoding plasmid were used to transfect the cells. GFP levels were monitored on Days 1 and 6 after transfection, by flow cytometry (Guava EasyCyte, Guava Technologies). Cell survival is expressed as a percentage and was calculated as a ratio: (meganuclease-transfected cell expressing GFP on Day 6/control transfected cell expressing GFP on Day 6), corrected for the transfection efficiency determined on Day 1.

γ–H2AX immunocytochemistry

For γ-H2AX immunocytochemistry, CHO-K1 cells were transfected with a mixture containing various amounts of plasmid encoding a HA-tagged meganuclease, with DNA levels made up to 4 µg with empty vector, in the presence of Polyfect reagent (Qiagen). Cells were fixed 48 h after transfection, by incubation with 2% of paraformaldehyde for 30 min, and permeabilized by incubation for 5 min at room temperature in 0.5% Triton. Cells were washed and incubated for 1 h in 0.3% Triton buffer supplemented with 10% normal goat serum (NGS) and 3% BSA, to block non-specific staining. Cells were then incubated for 1 h at RT with anti-γ-H2AX (Upstate: 1/10 000) and anti-HA (Santa Cruz:1/200) antibodies diluted in Triton 0.3% in PBS supplemented with 3% BSA and 10% NGS and then for 1 h with Alexa Fluor 488 goat anti-mouse (Invitrogen-Molecular Probes: 1/1000) and Alexa Fluor 546 goat anti-rabbit secondary antibodies diluted in 0.3% Triton in PBS supplemented with 3% BSA, and 10% NGS. Coverslips were incubated with 1 µg/ml 4′,6-diamino-2′-phenylindole (DAPI; Sigma), mounted and the γ-H2AX foci were visualized in transfected cells (HA-positive) by fluorescence microscopy. The inactive double-mutant V2(G19S)/V3(G19S) heterodimer was used as a negative control at a dose of 1 µg of each expression plasmid.

Endogenous gene targeting experiments

The donor plasmid for gene targeting experiments contained left and right homology arms generated by PCR amplification of the human *RAG1* locus. An exogenous DNA fragment was inserted between these two arms. These sequences consisted of either a 1.7-kb DNA fragment derived from a neomycin expression plasmid or a 10-bp DNA sequence containing an HindIII recognition site. The human 293H cells (Invitrogen) were plated at a density of 1×10^6 cells per 10 cm dish in complete medium (DMEM supplemented with 2 mM L-glutamine, penicillin

(100 IU/ml), streptomycin (100 μg/ml), amphotericin B (Fongizone: 0.25 μg/ml, Invitrogen-Life Science) and 10% FBS). The next day, cells were transfected in the presence of Lipofectamine 2000 transfection reagent (Invitrogen) according to the manufacturer's protocol. Cells were co-transfected with 2 μg of the donor plasmid and 3 μg of meganuclease expression vector. After 48 h of incubation at 37°C, cells were treated with trypsin and dispensed at a density of 10 or 100 cells per well in 96-well plates or re-plated and individual clones picked and subsequently amplified. DNA was extracted with the ZR-96 genomic DNA kit (Zymo research) according to the manufacturer's protocol. PCR amplification reactions were performed with the primers F2:5′-AGGATCT CCTGTCATCTCAC-3′ and R2: 5′-GCAGTGTTGCAG ATGTCACAG-3′, and F2:5′-AGGATCTCCTGTCATC TCAC-3′ and R12: 5′-CTTTCACAGTCCTGTACATCT TGT-3′ in order to detect the targeted integrations of the 10 bp and 1700 bp exogenous fragments, respectively. The PCR products related to the 10 bp integrations experiments were further digested by the restriction enzyme HindIII. Southern blot analysis was performed with genomic DNA digested with HindIII and hybridized with an 830 bp *RAG1* specific probe binding outside of the right homology arm of the donor plasmid.

RESULTS

Engineering of I-CreI derivatives cleaving a sequence from the human RAG1 gene

We described here an artificial heterodimeric meganuclease derived from the I-CreI meganuclease that is capable of cleaving a DNA sequence from the human *RAG1* gene, (12). We have further improved the cleavage activity of this protein by random mutagenesis, as previously described (32), resulting in the production of a new heterodimeric protein named, V2/V3. The V2 and V3 proteins have the following mutations with respect to the wild-type I-CreI sequence: N30R, Y33N, Q44A, R68Y, R70S and I77R for V2 and K28N, Y33S, Q38R, S40R, Q44Y, R70S, D75Q, I77V and E157G for V3. V2/V3 cleaves its cognate target efficiently in CHO-K1 cells (Figure 1a), in our previously described extrachromosomal assay based on meganuclease-induced recombination by single-strand annealing (SSA) (17). However, the formation of V2/V3 in target cells requires the simultaneous production of V2 and V3, which also results in the formation of V2/V2 and V3/V3 homodimers, cleaving the LL and RR targets, respectively (Figure 1a). These two homodimeric species are therefore likely to decrease overall specificity, by cleaving off target sequences. We used two approaches to abolish homodimer formation.

Design of obligate heterodimers

We first redesigned the dimerization interface to ensure that only heterodimers could be produced. We have previously described I-CreI derived obligate heterodimers functional *in vitro* (36). However, the extensive design resulted in a reduction of activity *in vivo* (data not shown). Here, we used a stepwise approach to preserve maximal activity. Several charged residues involved in electrostatic interactions between monomers, including Lys7, Glu8, Glu61, Lys96, Arg51 and Asp137, (Figure 1b) were mutated, as previously described (36). We also mutated the Gly19 residue, which has been shown to affect protein activity (32,33). In a stepwise approach, we first tested the impact of single mutations, such as K7E, E8K, E61R, K96E, R51D, D137R and G19S. The single mutations were introduced into the V2 and V3 mutants and the activity of the resulting homodimeric proteins against the two palindromic targets, LL or RR was monitored by simple transfection in CHO-K1 cells, using the SSA-based extrachromosomal assay described in Figure 1a. Supplementary Figure 1 shows that neither the K7E nor the E8K mutation significantly affected homodimer activity. In contrast, the E61R and K96E mutations greatly decreased homodimer activity for both V2 and V3 proteins, although they had a greater impact on the V2 protein than on the V3 protein. The G19S, R51D and D137R mutations almost entirely abolished cleavage activity.

We then co-expressed the various V2 mutants with a V3-derived partner bearing the compensatory mutation in CHO-K1 cells (For example the V2 variant carrying the K7E was associated with the V3 variant bearing the E8K mutation). The ability of each pair to cleave the LR, RR, or LL targets was monitored in our extrachromosamal assay in CHO-K1 cells (Supplementary Figure 2) and is expressed as a relative activity to the V2/V3 activity. V2 and V3 co-expression results in cleavage of the three LR, LL and RR targets as a result of the presence of the heterodimer and the two homodimers. The co-expression of V2/V3 variants with the R51D or D137R mutation results in very low levels of LR activity and were thus not further considered. The co-expression of V2 and V3 mutants carrying mutations at positions 61 and 96 leads to a large decrease in homodimer activities, particularly for the LL target, the cleavage of which was not detectable. A significant decrease of the activity against the LR target, corresponding to 60% of the V2/V3 activity, was also observed. When a V2 mutant carrying the K7E or E8K mutations was co-expressed with a V3 mutant carrying the compensatory mutations, heterodimer activity was not affected while homodimer activities were still relatively strong. It can be noted that in the case of co-expression, the LL cleavage activity due to the V2 E8K mutant was much weaker than the initial V2 activity, despite the V2 E8K mutant being clearly active against the LL palindromic target. Interestingly, the addition of the G19S mutation to V2 or V3 resulted in the largest decrease in the corresponding homodimer activity, with no significant decrease in the heterodimer activity (Supplementary Figures 1 and 2). However, we could not identify any symmetrical mutation that could be combined with G19S, in order to inactivate both homodimers, but preserved the heterodimer's activity.

With a view to obtaining true obligate heterodimers, the two sets of mutations, K7E-E8K and E61R-K96E, were combined with or without the G19S mutation. For both I-CreI-derived variants, V2 and V3, the four double mutants—K7E-E61R, E8K-E61R, K7E-K96E, E8K-K96E—were produced. The homodimer activity of the

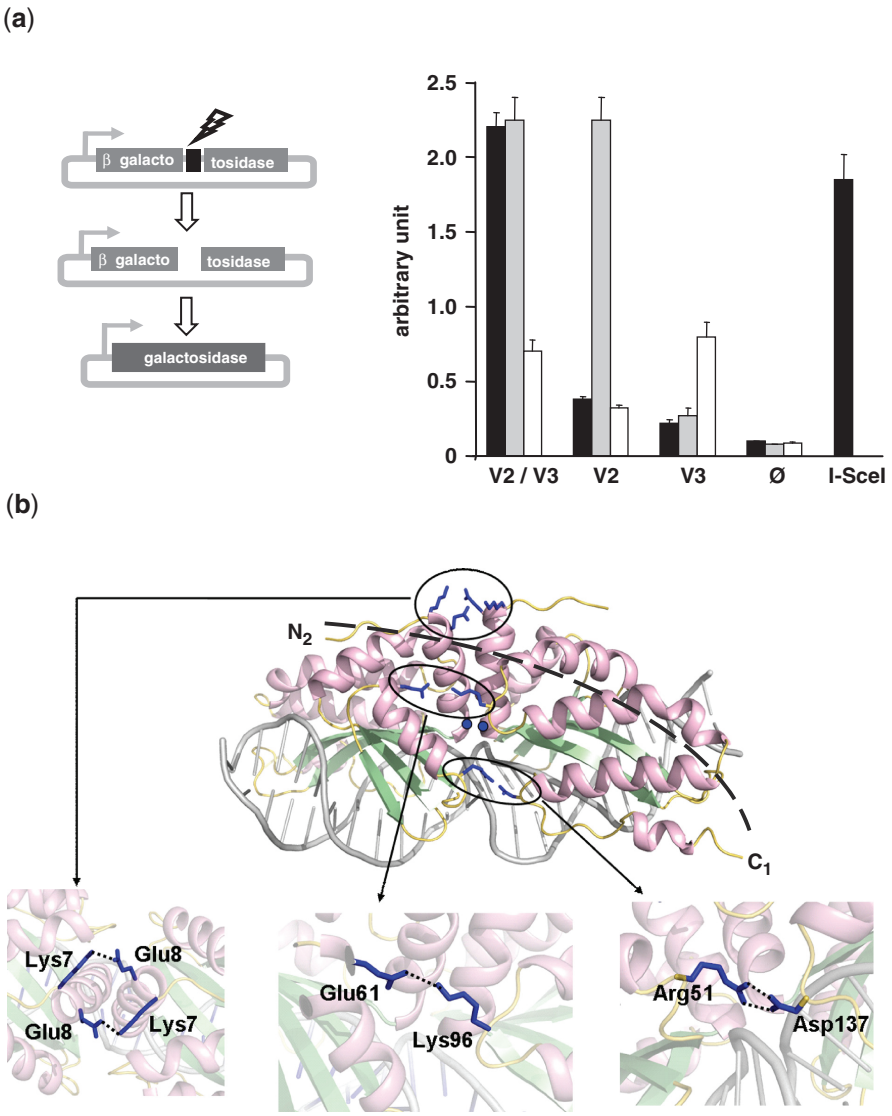

Figure 1. Design of obligate heterodimers and single-chain molecules. (**a**) Extrachromosomal assay used to monitor the activity of the engineered meganucleases in CHO-K1 cells. A meganuclease cleavage site (black box) is placed between two direct repeats in a reporter plasmid. Cleavage of the target site induces tandem-repeat recombination, thereby restoring a functional β-galactosidase gene. V2/V3: co-expression of the V2 and V3 mutant constructs; V2 and V3: expression of V2 and V3, respectively; Ø: empty vector. Cleavage activity was monitored against the palindromic targets LL (gray bars) and RR (white bars), and the non-palindromic target LR (black bars). The cleavage activity of I-SceI with its target is shown as a positive control (hatched bar). Bars indicate the mean ± SD (n = 3). (**b**) Engineered positions are indicated on the structure of the I-CreI meganuclease bound to its DNA substrate (PDB code 1G9Y). The three monomer–monomer interactions targeted in the obligate heterodimer design are highlighted. In addition, the position of the two Gly19 residues is indicated by two blue spheres. C_1 and N_2 indicate the N- and C-termini of the first and second monomer, respectively. The dashed gray line represents a linker used to generate single-chain molecules.

double mutants was monitored against the palindromic targets LL and RR in CHO-K1 cells in our extrachromosomal assay, after the individual transfection of each mutant. The activity of the four double mutants derived from V2 against the LL target was much lower than that of V2, almost reaching background levels (Figure 2a). The V3-derived mutants also displayed lower levels of activity, particularly those carrying the K96E mutation. Combinations of mutant proteins carrying compensatory mutations were then produced by co-expression in CHO-K1 cells and activity against the LR target was monitored (data not shown). All four active obligate heterodimers had similar cleavage activity and the meganuclease resulting from the co-expression of V2 E8K-E61R ($V2^+$) with V3 K7E-K96E ($V3^-$) was selected for further studies and characterization.

The abolition of homodimer activities may result from a lack of homodimer production or from the production of a non-functional protein. We addressed this question by determining the levels of different proteins in CHO-K1 cells. An HA-epitope was added to the C-terminus of the V2 and V2 E8K-E61R proteins and an S-Tag was added to the C-terminus of the V3 and V3 K7E-K96E proteins for this purpose. Protein levels were then monitored by western blotting after the transfection of CHO-K1 cells with a single mutant or with the two mutants

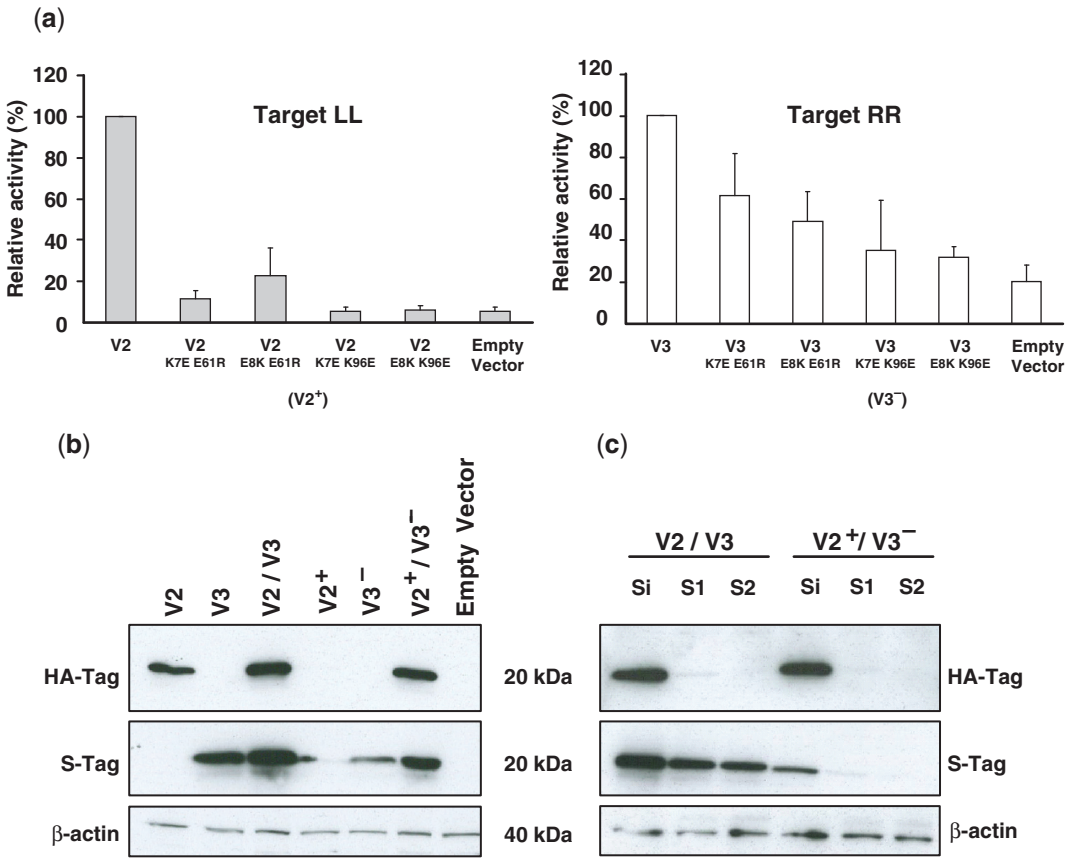

Figure 2. Activity of the double-mutant proteins derived from the V2 or V3 mutants. (**a**) Cleavage activity in an extrachromosomal assay in CHO-K1 cells (described in Figure 1a). Each homodimer and derivative was tested against its cognate palindromic target. V2 and V3: expression of the V2 and V3 constructs, respectively. Mutations are indicated beneath the protein. LL and RR targets; are the palindromic sequences cleaved by the V2/V2 and V3/V3 homodimeric proteins, respectively. The mutants chosen for further analysis are noted $V2^+$, $V3^-$. Bars indicate the mean ± SD ($n = 4$). (**b**) Expression profile of V2 and V3 derived mutants in CHO-K1 cells. Transfected CHO-K1 cells were harvested after 48 h and lysates were probed with antibodies against the HA-Tag or the S-Tag. V2/V3, $V2^+/V3^-$; co-expression of the V2 and V3, and $V2^+$ and $V3^-$ mutant constructs respectively. V2, $V2^+$, V3 and $V3^-$; expression of V2, $V2^+$, V3 and $V3^-$, respectively. The results are expressed as a percentage of the non-mutated protein activity (**c**). Depletion of HA-tagged proteins from the supernatant. The lysate of transfected CHO-K1 cells was depleted of HA-tagged proteins by two successive immunoprecipitations. The successive supernatants, S initial (Si), S1 and S2 were probed by anti-HA and anti-S-Tag antibodies on western blots. V2/V3, $V2^+/V3^-$; co-expression of the V2 and V3, and $V2^+$ and $V3^-$ mutant constructs respectively. Antibody against β-actin was used for the loading control.

that formed the V2/V3 or, the $V2^+/V3^-$ heterodimers. In co-expression experiments, the two mutants forming the V2/V3 heterodimer and the two proteins constituting the $V2^+/V3^-$ obligate heterodimer were produced in very similar amounts (Figure 2b). This similarity is consistent with the similarity in the activity of the two heterodimers against the LR target. The differences in the fate of individual proteins were much more striking. The V2 and V3 were produced in amounts equivalent to those observed for the V2/V3 heterodimer. By contrast, the $V2^+$ (V2 E8K-E61R) protein was not detected at all whereas the $V3^-$ (V3 K7E-K96E) protein level was produced in small amounts, corresponding to no more than one fifth the amount of V3 protein. Immunoprecipitation experiments were used to assess the presence or absence of the $V3^-$ homodimer when both partners of the $V2^+/V3^-$ heterodimer were expressed together. The lysates of CHO-K1 cells co-transfected either with the V2 and V3 coding vectors or with the $V2^+$ and $V3^-$ coding vectors were depleted of proteins carrying the HA epitope (V2 or $V2^+$) by successive incubations with an anti-HA antibody. The supernatants were analyzed by western blotting with anti-HA and anti-S-Tag antibodies (Figure 2c). After two successive immunoprecipitations, the two supernatants were entirely depleted of any HA-tagged proteins. In the case of the V2/V3 heterodimer, 50% of the S-Tag-labeled proteins (lane S2) were recovered after two successive immunoprecipitations, as shown by quantitative densitometry. This protein fraction corresponds to the V3/V3 homodimer. For the $V2^+/V3^-$ heterodimer, almost no $V3^-$ homodimer formation was detected, as revealed by the barely detectable signal after the first immunoprecipitation, showing that both $V2^+$ and $V3^-$ are required for stabilization and folding into a functional heterodimer.

Design of single-chain meganucleases

We and others have already described single-chain versions of I-CreI (37,38), an alternative approach to abolishing homodimer formation and activity. We designed a

Table 1. Amino acid sequences of the linker regions that were inserted between two I-CreI-variants V2 and V3

Number	Size (aa)	Primary sequence	Activity in yeast
1	22	$AA(GGGGS)_4$	+ +
2	35	AAGKSSDSKGIDLTNVTLPDTPTYSKAASDAIPPA	+ + +
3	30	AAGLEYPQAPYSSPPGPPCCSGSSGSSAGCS	+ + +
4	30	AAGLSYHYSNGGSPTSDGPALGGISDGGAT	+ + +
5	27	AAGDSSVSNSEHIAPLSLPSSPPSVGS	+ + +
6	25	AAGASQGCKPLALPELLTEDSYNTD	+ +
7	24	AAGNPIPGLDELGVGNSDAAAPGT	+ +
8	23	AAGAPTECSPSALTQPPSASGSL	+
9	27	$AA(GGGGS)_5$	+ +
10	17	$AA(GGGGS)_2$	+
11	12	$AA(GGGGS)_2$	+
12	17	AAGQVTSAAGPATVPSG	+
13	19	AAGGSPLKPSAPKIPIGGS	+
14	33	AAGGSPLKPSAPKIPIGGSPLKPSAPKIPIGGS	+ + +
15	15	AAGGSPLSKPIPGGS	+
16	25	AAGGSPLSKPIPGGSPLSKPIPGGS	+ +
17	31	AAGGSDKYNQALSERRAYVVANNLVSGGGGS	+ + +
18	32	AAGGSDKYNQALSKYNQALSKYNQALSGGGGS	+ + +

The activity of the resulting single-chain molecule was assessed in yeast. The linker selected for the RAG single-chain design is shown in underlined typeface.

series of new single-chain I-CreI derivatives, which maintain maximal levels of activity, by connecting the C-terminal part of a first monomer to the N-terminal part of a second monomer with a series of linkers differing in length and in their potential to form secondary structures (α-helix or type II polyproline helix). The activity of 18 such novel proteins was assessed (Table 1) in a yeast assay described in previous studies (17,37). A 32 amino acid linker with a predicted α-helix was eventually selected (number 18 in Table 1). In order to investigate the impact of the position of each monomer within the single-chain (sc) molecule and the impact of the G19S mutation, we produced the six constructs scV2-V3, scV2(G19S)-V3, scV2-V3(G19S), scV3-V2, scV3(G19S)-V2, scV3-V2(G19S) and assessed their ability to cleave the LL, RR, and LR targets (Supplementary Figure 3). All six molecules had high levels of activity against the LR target, equivalent to the activity of the V2/V3 heterodimer whereas cleavage of the LL and RR targets was either abolished or strongly decreased. These results demonstrate that the single-chain design favors intra-molecular interactions between the two I-CreI-derived moieties.

However, residual cleavage of the LL palindromic target was still observed for some constructs. We hypothesize that inter-molecular interactions could still occur, resulting in the generation of an active nuclease. The elimination of the LL cleavage activity could be achieved by the introduction of the G19S mutation into the V2 mutant or by the introduction of the engineered electrostatic interactions (Supplementary Figure 3). Finally, we combined the single-chain design and the new dimerization interface, to produce the $scV3^{-}$-$V2^{+}$ and $scV3^{-}$-$V2^{+}$ (G19S) molecules. These molecules were compared with previous constructs in an extrachromosomal *in vivo* assay (Figure 3a) or *in vitro* cleavage assay (Figure 3b). These two molecules had cleavage profiles very similar to those of their scV3-V2 and scV3-V2(G19S) counterparts, with G19S having again a strong impact on LR cleavage in CHO-K1 cells. The scV3-V2(G19S) and $scV3^{-}$-$V2^{+}$(G19S) molecules and the two obligate heterodimers $V2^{+}/V3^{-}$ and $V2^{+}$ (G19S)/$V3^{-}$, performed similarly, displaying high levels of cleavage activity and no detectable activity against the palindromic targets.

Obligate and single-chain meganucleases are active in a gene targeting assay in CHO-K1 cells

We compared the abilities of different V2/V3 variants to stimulate homologous gene targeting (GT) in a previously described chromosomal reporter system (32), in CHO-K1 cells bearing a single chromosomal copy of the *LacZ* gene interrupted by the LR target (Figure 4a). GT frequency was monitored over a wide dose range. The V2/V3 heterodimer induced GT with a frequency of up to 0.27%. The G19S, obligate heterodimer, and G19S obligate heterodimer versions increased this frequency up to 0.8% (Figure 4a). However, the frequency of GT observed with these molecules rapidly decreased at higher doses, suggesting potential adverse effects. By contrast, the two single-chain molecules were the most similar in behavior to the I-SceI meganuclease, and were highly effective over a broad range of doses, their activity level reaching a near plateau at about 100 ng of expression vector. Western blotting analysis demonstrated that this plateau was not due to an expression or stability artifact, for all the tested meganucleases were produced in similar amounts (Figure 4b and data not shown). We also failed to detect any truncated products (Figure 4b and data not shown), showing that the linker of the single-chain molecules was not subjected to proteolytic cleavage in CHO-K1 cells.

Obligate heterodimer and single-chain designs decrease toxicity and off-site cleavage

Toxicity is a major issue for DSB-induced recombination technology, particularly for therapeutic applications,

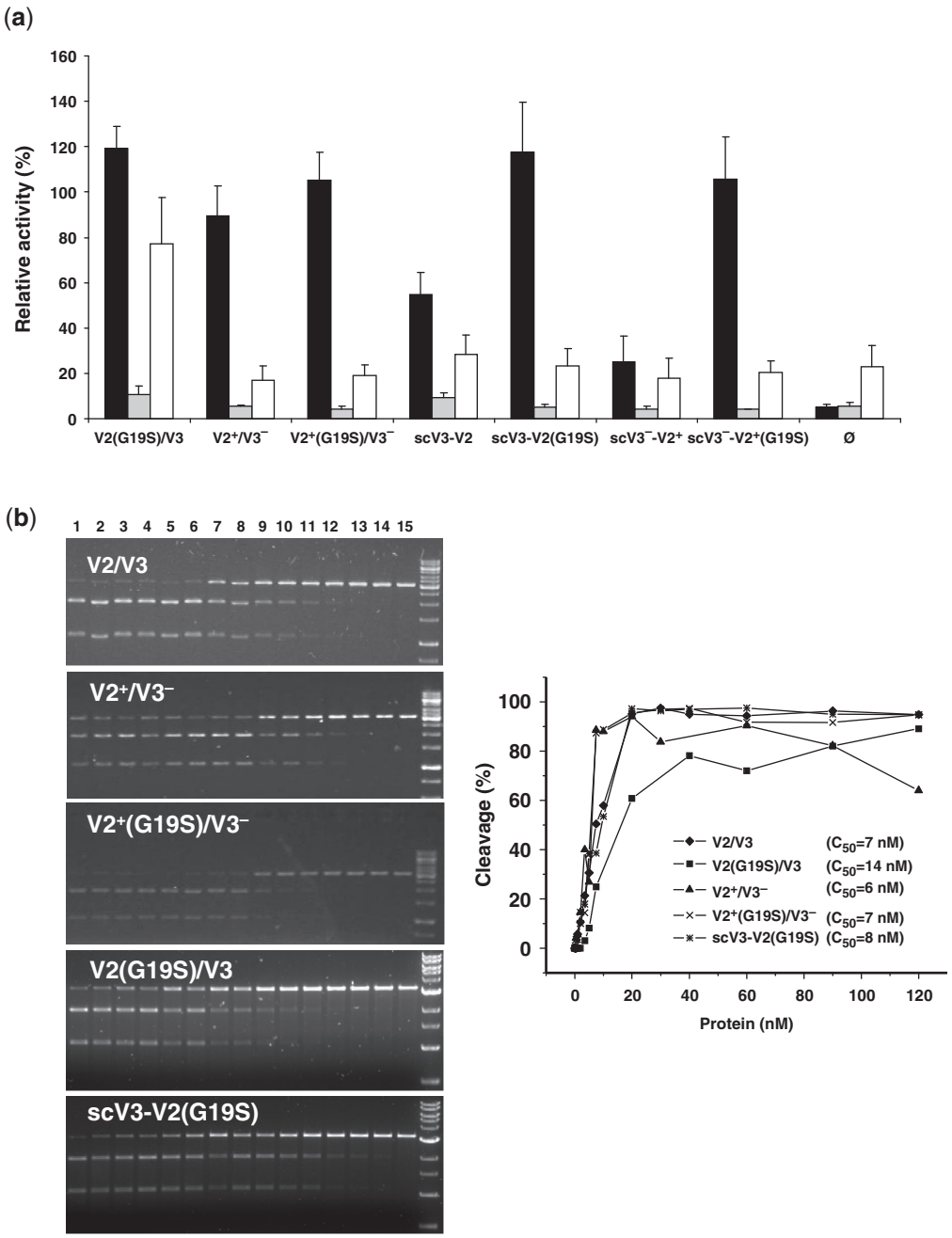

Figure 3. Activity of the obligate heterodimers and single-chain meganucleases (**a**) Extrachromosomal assay in CHO-K1 cells. The principle of the assay is described in Figure 1a. Cells were transfected with meganuclease constructs of various designs and the resulting cleavage of the three targets LR (black bars), LL (gray bars) and RR (white bars) is shown as a percentage of the activity resulting from V2/V3 co-expression. V2+, mutant V2 (E8K, E61R); V3−, mutant V3 (K7E, K96E); sc, single-chain molecule; Ø: empty vector. Bars indicate the mean ± SD ($n = 3$). (**b**) *In vitro* cleavage activity of heterodimers and single-chain molecules against the LR target. The reaction mixture included a target concentration of 2 nM and the purified protein. Gels from the left panel were quantified by densitometry. The graph in the right panel summarizes the data. Lanes 1–15; protein concentrations of 120, 90, 60, 40, 30, 20, 10, 7.5, 5, 3.5, 2, 1, 0.5, 0.25 and 0 nM, respectively. C_{50}; enzyme concentration required to cleave 50% of the target DNA.

for which the activity/toxicity ratio is of major concern. The toxicity of the RAG1 meganucleases was evaluated by adapting a previously described cell survival test (11) to the CHO-K1 cells used for the activity dose response assay. The V2(G19S)/V3(G19S) heterodimer was used as a control for non-toxicity, as this molecule was completely inactive in all our assays (Figure 4a and data not shown). At the active dose (~100 ng, the dose at which the meganucleases displayed their maximum level of activity Figure 4b), toxicity was barely detectable with all meganucleases (Figure 5a). However, the V2/V3 heterodimer displayed significant toxicity at higher doses, which was partly alleviated by the obligate heterodimer design (but not by the G19S mutation), and further decreased by the

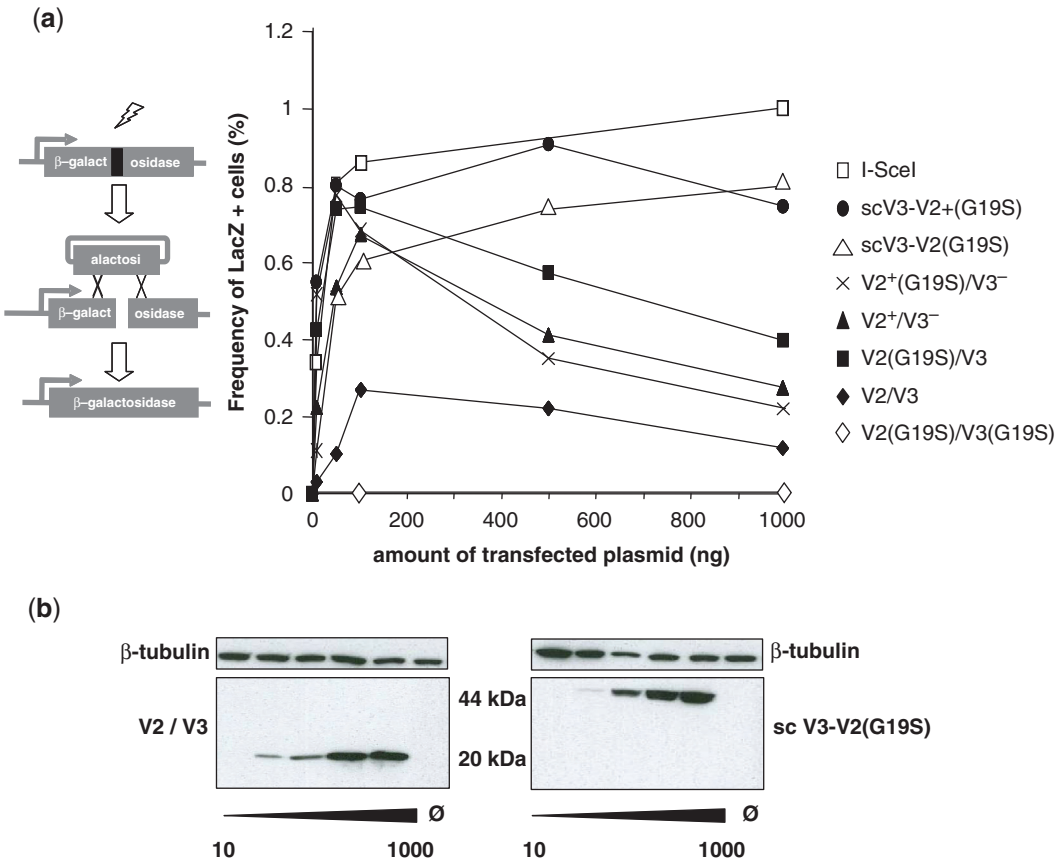

Figure 4. Gene targeting in a chromosomal reporter system in CHO-K1 cells. (**a**) Chromosomal assay used to monitor gene targeting induced by engineered meganucleases in CHO-K1 cells. Left panel: CHO-K1 cells containing a single-chromosomal copy of the *lacZ* gene interrupted by the DNA target site (black box) were co-transfected with meganuclease-expressing vectors and a repair plasmid. Target cleavage induces a gene targeting event that restores a functional β-galactosidase gene, by homologous recombination between the chromosomal reporter and the repair plasmid. We monitored the activity of the RAG1 meganucleases and of I-SceI, using two different cell lines, differing only in the meganuclease target site (LR for the RAG1 meganucleases, I-SceI target for the I-SceI meganuclease). Right panel: dose-response study. CHO-K1 cell lines were transfected with various amounts of expression vector for various meganucleases and a fixed quantity of the repair plasmid. (**b**) Meganuclease levels in CHO-K1 cells. Protein levels were monitored in CHO-K1 cells by western blotting, using the same amounts of plasmid as for the activity study. V2/V3, co-expression of the V2 and V3 constructs. scV3-V2(G19S), production of the single-chain protein. Protein levels were analyzed after transfection with various quantities of expression plasmids, from 10 to 1000 ng. Ø: empty vector.

single-chain designs, with the best version closely mimicking the pattern obtained with I-SceI, the most specific endonuclease to date.

The toxicity of sequence-specific endonucleases is usually attributed to off-site cleavage, which can result in mutations, deletions, translocations and other gross genomic alterations (9). We evaluated off-site cleavage, by monitoring the formation of phosphorylated H2AX histone (γ-H2AX) foci. γ-H2AX focus formation is one of the first responses of the cell to DNA DSBs and provides a convenient means of monitoring DSBs in living cells (40). CHO-K1 cells were transfected with two doses of meganuclease expression vectors (active dose and 10-fold excess) and γ-H2AX was monitored by image analysis (Figure 5b). At the active dose, with the exception of V2 (G19S)/V3, the expression of the various RAG meganucleases induced between 1 and 5 γ-H2AX foci per transfected cell, as did I-SceI. For discrimination between the different designs, we analyzed the induction of γ-H2AX foci when meganucleases were produced in large excess (10-fold excess). The production of significant numbers of foci was observed for the V2/V3 and V2 (G19S)/V3 heterodimers at an excess dose (Figure 5b), with up to 70% of the transfected cells showing more than 25 γ-H2AX foci per cell. However, the obligate heterodimer and single-chain designs drastically reduced this effect, close to the level found with I-SceI and the V2(G19S)/V3(G19S) proteins.

A single-chain meganuclease induces high levels of homologous gene targeting at the endogenous human *RAG1* locus

Finally, we evaluated the ability of the scV3^{-}-V2^{+}(G19S) engineered meganuclease to stimulate homologous recombination at the endogenous *RAG1* locus in human cells. We designed two repair plasmids. The first contained 10 bp of exogenous DNA (containing a single HindIII restriction site) flanked by two sequences, 900 bp and 500 bp in length, homologous to the human *RAG1* locus. The second carried 1700 bp of exogenous DNA flanked

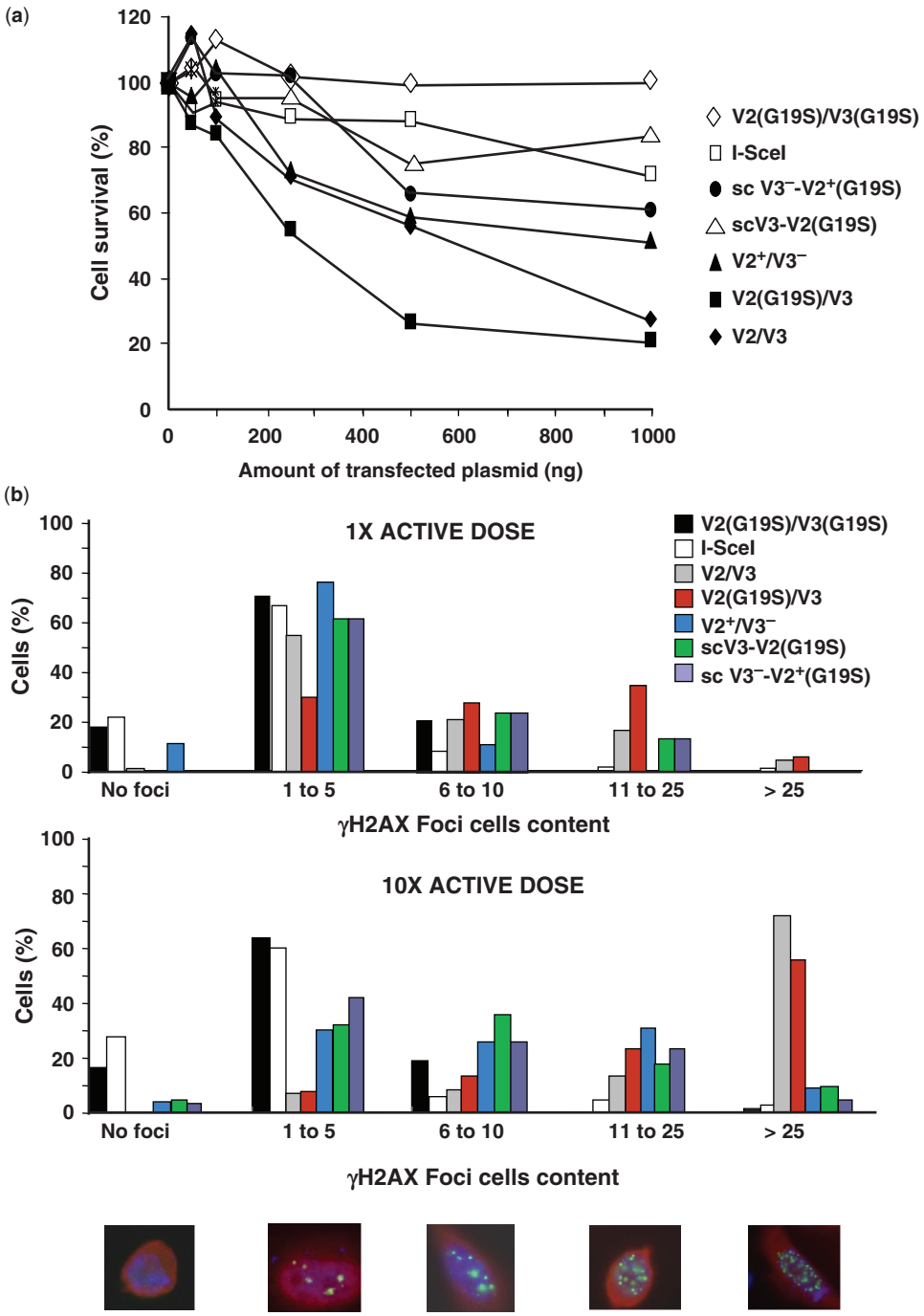

Figure 5. Evaluation of the toxicity of meganucleases. (**a**) Toxicity of the engineered meganucleases, as monitored by a cell survival assay. Various amounts of meganuclease expression vector and a constant amount of plasmid encoding GFP were used to co-transfect CHO-K1 cells. Cell survival is expressed as the percentage of cells expressing GFP 6 days after transfection, as described in the 'Materials and Methods' section. The inactive V2(G19S)/V3(G19S) heterodimer is shown as a control for non-toxicity. (**b**) DNA damage was visualized by the formation of γ-H2AX foci at DNA double-strand breaks. CHO-K1 cells were transfected with a plasmid encoding the meganuclease fused to an HA epitope at the active dose (quantity used for maximal activity) or 10 times the active dose. The inactive V2(G19S)/V3(G19S) heterodimer was used as a control (1 μg of each expression plasmid). The number of γ-H2AX foci in transfected cells (HA positive) was determined for each condition. Representative images below each class illustrate the different classes of γ-H2AX foci in CHO-K1 cells transfected with meganuclease: Red, HA labeling; green, γ-H2AX foci; and blue, DAPI staining.

by two homology arms of 2.0 kb and 1.6 kb in length. We assessed the ability of the scV3^{-}-V2^{+}(G19S) protein to induce targeted events at the endogenous locus (Figure 6). After co-transfection with the meganuclease-encoding vector and the donor repair plasmid, 293H cells were cloned and cultured in 96-well microplates and analyzed by PCR. PCR screening is exemplified in Figure 6b and c and summarized in Table 2. Positive clones were

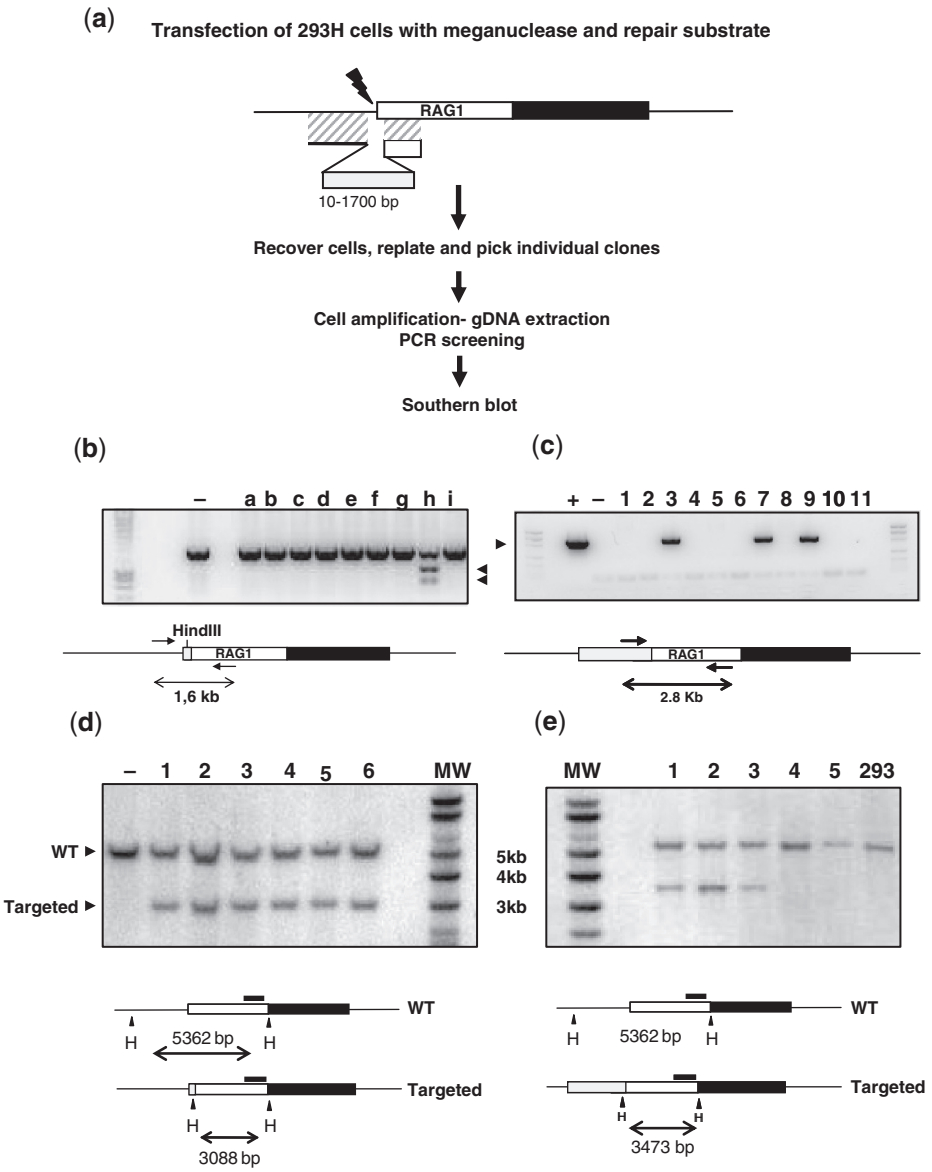

Figure 6. Targeted integration at the endogenous *RAG1* locus, driven by a single-chain meganuclease. (**a**) Experimental outline and diagram of the gene targeting strategy used at the endogenous *RAG1* locus. The *RAG1* target sequence is located just upstream from the coding sequence for exon 2 of the Rag1 protein. Exon 2 is boxed, with the open reading frame in white. Cleavage of the native *RAG1* gene by the meganuclease yields a substrate for homologous recombination, which may use the repair plasmid containing either 10 bp or 1.7 kb of exogenous DNA. The 10-bp DNA fragment (containing a single HindIII restriction site) is flanked by two sequences, 900 bp and 500 bp in length, homologous to the human *RAG1* locus. The 1700 bp DNA fragment is flanked by two homology arms of 2.0 kb and 1.6 kb in length. Human 293H cells were transfected with 3 μg of scV3^{-}-V2^{+}(G19S) meganuclease expression plasmid and 2 μg of the repair substrate, cultured without selection for 48 h, re-plated, and individual clones picked and amplified. Targeted integration events were detected by PCR amplification and subsequently confirmed by Southern blotting. (**b**) Example of a screen for targeted integration events with a repair plasmid containing 10 bp of exogenous DNA. Genomic DNA derived from individual clones (a–i) obtained after transfection with a meganuclease expression plasmid and a repair plasmid was amplified by PCR and digested with HindIII. − is the negative control (non-transfected cells). PCR primers positions are indicated by simple arrows, with one primer binding outside the region of homology used in the repair plasmid. On the gel, bands correlated with a targeted event are indicated with a black triangle (**c**) Example of a PCR screen for targeted integration events with the repair plasmid containing 1700 bp of exogenous DNA. PCR analysis of 11 different samples (1–11), + indicates positive control (a recombinant plasmid), and − is the negative control (non-transfected cells). (**d**) Southern blot analysis of six individual clones (1–6) that were positive in the PCR screen for the integration of the 10 bp exogenous fragment, analyzed together with DNA from non-transfected cells (−). The locus maps indicate the restriction pattern of the wild-type locus (5.3 kb) and the targeted locus (3.0 kb). (**e**) Southern blot analysis of three PCR-positive (1–3) and two PCR-negative clones (4–5) samples derived from single transfected cells are analyzed, together with DNA from non-transfected cells (−). The locus maps indicate the restriction pattern of the wild-type locus (5.3 kb) and the targeted locus (3.4 kb). Probes are indicated by a solid black box. H, genomic location of HindIII cleavage sites.

amplified, double-checked by PCR, and analyzed by Southern blotting. The insertion of the 10 bp and 1700 bp fragments was confirmed by Southern blotting in 93% (13/14) and 82% (9/11) of the initial PCR positive clones, respectively (see examples in Figure 6d and e). Three clones were not confirmed by Southern blot or by PCR after amplification, and were considered as false positives. We hypothesize that this is due to the presence

Table 2. Gene targeting experiments performed with scV3^{-}-V2^{+}(G19S) engineered to target the human *RAG1* locus

Insertion length (bp)	Meganuclease	Cells/ well	Total no. of cells analyzed	PCR Positives	Confirmed by Southern blot confirmed/analyzed	Recombination frequency(%)
10	+	1	183	7 (3.8%)	6/6	3.8%
10	+	1	376	13 (3.5%)	7/8	3.0%
10	–	1	325	0		
10	–	1	257	0		
1700	+	1	94	5 (5.3%)	3/5	3.2%
1700	+	1	376	23 (6.1%)	6/6	6.1%
1700	–	1	94	0		
1700	–	1	351	0		
1700	–	10	2560	0		
1700	–	10	1880	0		
1700	–	100	37 000	0		

Each line corresponds to the results of an independent experiment. Cells/well: number of cells used in the study of clones (1 cell/well) or pools (10 or 100 cells/well) of transfected cells. Recombination frequencies are calculated as the frequency of PCR positives, corrected by the ratio of clones confirmed by Southern blotting.

of a mixed colony resulting from a contamination of adjacent cells during the picking of individual clones, followed by loss of the positive clone during amplification. Nevertheless, meganuclease-mediated gene insertion could be achieved in up to 6% of the transfected cells (Table 2). In contrast, no GT event was detected in the absence of meganuclease, with both repair matrices, out of a total of 582 (10 bp insertion matrix) and 445 (1700 bp insertion matrix) clones.

In order to screen a larger number of cells transfected with the repair matrix alone (1700 bp insertion matrix), we also analyzed pools of transfected cells. Our PCR test was first assessed for its ability to detect positives in non-clonal populations: targeted cells displaying the 1700 bp insertion were diluted in naive 293H cells and analyzed by PCR to test for the presence of the insertion. Targeted cells could still be detected even when diluted in a 1000-fold excess of non-targeted cells (data not shown). Then, 96-well plates were seeded with cells transfected with the repair matrix alone, at density of 10 or 100 cells per well, and each pool was analyzed by PCR. Again, no positives were identified in the absence of meganuclease among more than 40 000 analyzed cells (Table 2). These results demonstrate that gene targeting can be achieved with a high efficiency and without selection at an endogenous locus in mammalian cells using an engineered meganuclease.

DISCUSSION

One of the most exciting prospects for meganuclease-induced gene targeting lies in its potential for use in molecular medicine, although such applications are highly demanding in terms of safety. The use of homodimeric scaffolds, such as I-CreI, to produce a heterodimeric meganuclease that recognizes and cleaves an asymmetric DNA sequence requires the engineering of two different monomers. Heterodimer formation is then obtained by producing the two monomers simultaneously in the cell. However, this leads to the presence of a mixture of enzymes within the cells and the two homodimer by-products could favor off-site cleavage jeopardizing overall meganuclease specificity. Two different strategies have been used to tackle this problem. The first one consists in the modification of the dimerization interface, to produce obligate heterodimers. Such an approach has previously been applied to zinc-finger nucleases, (34,35), resulting in improvements in specificity, and to I-CreI-derived obligatory heterodimers (36). Natural evolution has selected an alternative solution, single-chain meganucleases (41,42) such as I-SceI, which is considered to be the gold standard in terms of specificity. Artificial single-chain molecules derived from wild-type homodimeric LAGLIDADG meganucleases have already been described by us and others (37,38).

Extensive redesign can result in loss of activity. In this study, we tried to avoid any trade off between activity and specificity, by simplifying the designs. A previously described I-CreI-derived obligate heterodimer (36) displayed mutations at positions Lys7, Asp8, Glu61 and Lys96, but also in a hydrophobic patch made by several residues. In the new obligate heterodimers, we engineered only two electrostatic interactions (Lys7-Asp8 and Glu61-Lys96) that proved sufficient to destabilise homodimer formation, without any loss of activity for the heterodimer. For the single-chain designs, we previously connected two I-CreI monomers with a short linker (37), but the use of this shorter sequence was only made possible by truncating the first I-CreI monomer by one third. Here, we favored the use of long linkers, allowing for minimal modification of both monomers, as described recently (38). In addition, we could create a functional single-chain molecule carrying the obligate heterodimer specific mutations, a design that more closely mimics natural single-chain meganucleases, whose diverged LAGLIDADG regions are able to interact together in a 'heterodimer-like' manner, but can no longer form homodimers (22).

A G19S mutation had a striking effect on both dimers and single-chain molecules. This mutation was previously found to decrease the activity of dimers when it was

present in the two monomers, but to increase this activity when present only in one monomer (our unpublished data). Consistently, G19S strongly impaired or abolished the activity of the V2(G19S)/V2(G19S), V3(G19S)/V3(G19S) and V2(G19S)/V3(G19S) molecules (Supplementary Figure 2 and Figure 4). The Gly19 residue is found at the end of the LAGLIDADG helix, close to the active site, and its change for a Serine substitutes a hydrogen atom by a more voluminous side chain with a polar hydroxyl group. Thus, the inhibition of G19S/G19S dimer activity could be explained by a rearrangement of the interhelical packing that disrupts the catalytic center. In the scV3-V2(G19S) molecule, G19S could further impair cleavage of the LL target, possibly by alleviating any possible functional interactions between V2 moieties from different molecules. Activity enhancement is more difficult to explain. A similar effect had been observed previously with a heterodimer carrying a G19A mutation (32). In this study, G19S enhanced the activity of the V2/V3 heterodimer in the GT assay in CHO-K1 cells (Figure 4), but not *in vitro* (Figure 3b), suggesting that the mutation does not directly impact the specific activity of the protein. Improvement of the scV3-V2 by this sole mutation was also spectacular [compare scV3-V2 and scV3-V2(G19S) in Figure 3]. Further characterization will be necessary to understand the impact of the G19S mutation.

In their final versions, both obligate heterodimer and single-chain designs displayed high levels of activity, lower toxicity and minimal numbers of off-site cleavage events. Even at high doses (10-fold the dose required for optimal activity), these designs largely prevented the toxicity and γ–H2AX foci formation observed with both the V2/V3 and V2 (G19S)/V3 heterodimers. The single-chain designs (with or without obligate heterodimer mutations) appeared significantly superior to the obligate heterodimer in the GT assay in CHO-K1 cells, wherein single-chain molecules displayed high activity over a much broader dose range. It is difficult to explain these differences: whereas the obligatory heterodimer displayed a slightly higher toxicity than the single chain in the cell survival assay (Figure 5a), the γ-H2AX foci assay did not show any difference between the two designs (Figure 5b). These discrepancies could illustrate the limitations of the γ-H2AX foci assay, in which high levels of background observed in immortalized cell lines could hide significant differences between endonucleases.

These differences need to be confirmed with other engineered proteins cleaving different targets. Nevertheless, single-chain molecules have the major advantage of alleviating the problems associated with the control of co-expression. Promising results have also been obtained for the characterization of new scaffolds derived from natural monomeric endonucleases, such as I-DmoI (22,30,33) and I-AniI (43,44), hybrid I-DmoI/I-CreI proteins (37,45), and an hybrid of the *S. cerevisiae* PI-SceI protein and its *Candida tropicalis* homologue (18), and single-chain meganucleases, natural or artificial, represent a new generation of engineered endonucleases for genome engineering. In the present study, the properties of the RAG1 single-chain meganucleases place these molecules among the best molecular scissors available for genome surgery strategies. Engineered meganucleases have been used to target chromosomal reporter genes in mammalian cells (32,33), but this is the first time that the engineering of an endogenous gene by such proteins has been reported. Gene targeting was observed in 3–6% of total transfected cells. These endonucleases should facilitate gene correction therapy for monogenetic diseases, such as SCID.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Direction Générale des Entreprises du Ministère de l'Industrie et des Finances [convention no. 05290604]; EU Sixth Framework Programme for Research [contract LSHG-CT-2006-037226, MEGATOOLS]. Funding for open access charge: EU Sixth Framework Programme for Research (contract LSHG-CT-2006-037226, MEGATOOLS).

Conflict of interest statement. None declared.

REFERENCES

1. Jacquier,A. and Dujon,B. (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*, **41**, 383–94.
2. Choulika,A., Perrin,A., Dujon,B. and Nicolas,J.F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol. Cell Biol.*, **15**, 1968–1973.
3. Rouet,P., Smih,F. and Jasin,M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell Biol.*, **14**, 8096–8106.
4. Paques,F. and Duchateau,P. (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr. Gene. Ther.*, **7**, 49–66.
5. Pingoud,A. and Silva,G.H. (2007) Precision genome surgery. *Nat. Biotechnol.*, **25**, 743–744.
6. Cai,C.Q., Doyon,Y., Ainley,W.M., Miller,J.C., Dekelver,R.C., Moehle,E.A., Rock,J.M., Lee,Y.L., Garrison,R., Schulenberg,L. *et al.* (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol. Biol.*, **69**, 699–709.
7. Doyon,Y., McCammon,J.M., Miller,J.C., Faraji,F., Ngo,C., Katibah,G.E., Amora,R., Hocking,T.D., Zhang,L., Rebar,E.J. *et al.* (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat. Biotechnol.*, **26**, 702–708.
8. Lombardo,A., Genovese,P., Beausejour,C.M., Colleoni,S., Lee,Y.L., Kim,K.A., Ando,D., Urnov,F.D., Galli,C., Gregory,P.D. *et al.* (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.*, **25**, 1298–1306.
9. Porteus,M.H. and Carroll,D. (2005) Gene targeting using zinc finger nucleases. *Nat. Biotechnol.*, **23**, 967–973.
10. Urnov,F.D., Miller,J.C., Lee,Y.L., Beausejour,C.M., Rock,J.M., Augustus,S., Jamieson,A.C., Porteus,M.H., Gregory,P.D. *et al.* (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, **435**, 646–651.
11. Maeder,M.L., Thibodeau-Beganny,S., Osiak,A., Wright,D.A., Anthony,R.M., Eichtinger,M., Jiang,T., Foley,J.E., Winfrey,R.J., Townsend,J.A. *et al.* (2008) Rapid 'open-source' engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell*, **31**, 294–301.

12. Bibikova,M., Beumer,K., Trautman,J.K. and Carroll,D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764.
13. Cannata,F., Brunet,E., Perrouault,L., Roig,V., Ait-Si-Ali,S., Asseline,U., Concordet,J.P. and Giovannangeli,C. (2008) Triplex-forming oligonucleotide-orthophenanthroline conjugates for efficient targeted genome modification. *Proc. Natl Acad. Sci. USA*, **105**, 9576–9581.
14. Eisenschmidt,K., Lanio,T., Simoncsits,A., Jeltsch,A., Pingoud,V., Wende,W. and Pingoud,A. (2005) Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Res.*, **33**, 7039–7047.
15. Smith,J., Grizot,S., Arnould,S., Duclert,A., Epinat,J.-C., Prieto,J., Redondo,P., Blanco,F.J., Bravo,J., Montoya,G. *et al.* (2006) A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.*, **34**, e149.
16. Ashworth,J., Havranek,J.J., Duarte,C.M., Sussman,D., Monnat,R.J. Jr., Stoddard,B.L. and Baker,D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature*, **441**, 656–659.
17. Arnould,S., Chames,P., Perez,C., Lacroix,E., Duclert,A., Epinat,J.C., Stricher,F., Petit,A.S., Patin,A., Guillier,S. *et al.* (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J. Mol. Biol.*, **355**, 443–458.
18. Steuer,S., Pingoud,V., Pingoud,A. and Wende,W. (2004) Chimeras of the homing endonuclease PI-SceI and the homologous Candida tropicalis intein: a study to explore the possibility of exchanging DNA-binding modules to obtain highly specific endonucleases with altered specificity. *Chembiochem*, **5**, 206–213.
19. Eastberg,J.H., McConnell Smith,A., Zhao,L., Ashworth,J., Shen,B.W. and Stoddard,B.L. (2007) Thermodynamics of DNA target site recognition by homing endonucleases. *Nucleic Acids Res.*, **35**, 7209–7221.
20. Niu,Y., Tenney,K., Li,H. and Gimble,F.S. (2008) Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J. Mol. Biol.*, **382**, 188–202.
21. Rosen,L.E., Morrison,H.A., Masri,S., Brown,M.J., Springstubb,B., Sussman,D., Stoddard,B.L. and Seligman,L.M. (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.*, **34**, 4791–4800.
22. Silva,G.H., Belfort,M., Wende,W. and Pingoud,A. (2006) From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes. *J. Mol. Biol.*, **361**, 744–754.
23. Chen,Z. and Zhao,H. (2005) A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res.*, **33**, e154.
24. Stoddard,B.L. (2005) Homing endonuclease structure and function. *Q Rev Biophys*, **38**, 49–95.
25. Dalgaard,J.Z., Garrett,R.A. and Belfort,M. (1993) A site-specific endonuclease encoded by a typical archaeal intron. *Proc. Natl Acad. Sci. USA*, **90**, 5414–5417.
26. Silva,G.H., Dalgaard,J.Z., Belfort,M. and Van Roey,P. (1999) Crystal structure of the thermostable archaeal intron-encoded endonuclease I-DmoI. *J. Mol. Biol.*, **286**, 1123–1136.
27. Chevalier,B.S., Monnat,R.J. Jr. and Stoddard,B.L. (2001) The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat. Struct. Biol.*, **8**, 312–316.
28. Chevalier,B., Turmel,M., Lemieux,C., Monnat,R.J. Jr. and Stoddard,B.L. (2003) Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J. Mol. Biol.*, **329**, 253–269.
29. Moure,C.M., Gimble,F.S. and Quiocho,F.A. (2003) The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J. Mol. Biol.*, **334**, 685–695.
30. Marcaida,M.J., Prieto,J., Redondo,P., Nadra,A.D., Alibes,A., Serrano,L., Grizot,S., Duchateau,P., Paques,F., Blanco,F.J. *et al.* (2008) Crystal structure of I-DmoI in complex with its target DNA provides new insights into meganuclease engineering. *Proc. Natl Acad. Sci. USA*, **105**, 16888–16893.
31. Chevalier,B., Sussman,D., Otis,C., Noel,A.J., Turmel,M., Lemieux,C., Stephens,K., Monnat,R.J. Jr. and Stoddard,B.L. (2004) Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry*, **43**, 14015–14026.
32. Arnould,S., Perez,C., Cabaniols,J.P., Smith,J., Gouble,A., Grizot,S., Epinat,J.C., Duclert,A., Duchateau,P. and Paques,F. (2007) Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J. Mol. Biol.*, **371**, 49–65.
33. Redondo,P., Prieto,J., Munoz,I.G., Alibes,A., Stricher,F., Serrano,L., Cabaniols,J.P., Daboussi,F., Arnould,S., Perez,C. *et al.* (2008) Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature*, **456**, 107–111.
34. Miller,J.C., Holmes,M.C., Wang,J., Guschin,D.Y., Lee,Y.L., Rupniewski,I., Beausejour,C.M., Waite,A.J., Wang,N.S., Kim,K.A. *et al.* (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.*, **25**, 778–785.
35. Szczepek,M., Brondani,V., Buchel,J., Serrano,L., Segal,D.J. and Cathomen,T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.*, **25**, 786–793.
36. Fajardo-Sanchez,E., Stricher,F., Paques,F., Isalan,M. and Serrano,L. (2008) Computer design of obligate heterodimer meganucleases allows efficient cutting of custom DNA sequences. *Nucleic Acids Res.*, **36**, 2163–2173.
37. Epinat,J.C., Arnould,S., Chames,P., Rochaix,P., Desfontaines,D., Puzin,C., Patin,A., Zanghellini,A., Paques,F. and Lacroix,E. (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.*, **31**, 2952–2962.
38. Li,H., Pellenz,S., Ulge,U., Stoddard,B.L. and Monnat,R.J. Jr. (2009) Generation of single-chain LAGLIDADG homing endonucleases from native homodimeric precursor proteins. *Nucleic Acids Res.*, **37**, 1650–1662.
39. Mombaerts,P., Iacomini,J., Johnson,R.S., Herrup,K., Tonegawa,S. and Papaioannou,V.E. (1992) RAG-1-deficient mice have no mature B and T lymphocytes. *Cell*, **68**, 869–877.
40. Rogakou,E.P., Pilch,D.R., Orr,A.H., Ivanova,V.S. and Bonner,W.M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.*, **273**, 5858–5868.
41. Posey,K.L., Koufopanou,V., Burt,A. and Gimble,F.S. (2004) Evolution of divergent DNA recognition specificities in VDE homing endonucleases from two yeast species. *Nucleic Acids Res.*, **32**, 3947–3956.
42. Chevalier,B.S. and Stoddard,B.L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.*, **29**, 3757–3774.
43. Takeuchi,R., Certo,M., Caprara,M.G., Scharenberg,A.M. and Stoddard,B.L. (2009) Optimization of in vivo activity of a bifunctional homing endonuclease and maturase reverses evolutionary degradation. *Nucleic Acids Res.*, **37**, 877–890.
44. McConnell Smith,A., Takeuchi,R., Pellenz,S., Davis,L., Maizels,N., Monnat,R.J. Jr. and Stoddard,B.L. (2009) Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl Acad. Sci. USA*, **106**, 5099–5104.
45. Chevalier,B.S., Kortemme,T., Chadsey,M.S., Baker,D., Monnat,R.J. and Stoddard,B.L. (2002) Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell*, **10**, 895–905.