(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 August 2004 (12.08.2004)

PCT

(10) International Publication Number
WO 2004/067736 A2

(51) International Patent Classification[7]: C12N 9/22, 15/55, 15/10, A61K 48/00, A01K 67/027, A01H 1/00

(21) International Application Number:
PCT/IB2004/000827

(22) International Filing Date: 28 January 2004 (28.01.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/442,911	28 January 2003 (28.01.2003)	US
60/491,535	1 August 2003 (01.08.2003)	US

(71) Applicant *(for all designated States except US)*: **CELLECTIS** [FR/FR]; 102 route de Noisy, F-93230 Romainville (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ARNOULD, Sylvain** [FR/FR]; 116 rue Lamarck, F-75018 Paris (FR). **BRUNEAU, Sylvia** [FR/FR]; 30 rue des Favorites, F-75015 Paris (FR). **CABANIOLS, Jean-Pierre** [FR/FR]; 17 rue des Eglantines, F-95320 Saint Leu La Forêt (FR). **CHAMES, Patrick** [FR/FR]; 123 rue du Château, F-75014 Paris (FR). **CHOULIKA, André** [FR/FR]; 164 rue de la Croix Nivert, F-75015 Paris (FR). **DUCHATEAU, Philippe** [FR/FR]; 17 rue Michel Cadoret, F-91590 Cerny (FR). **EPINAT, Jean-Charles** [FR/FR]; 20 allée Georges Récipon, F-75019 Paris (FR). **GOUBLE, Agnès** [FR/FR]; 26 rue Saint-Paul, F-75004 Paris (FR). **LACROIX, Emmanuel** [BE/FR]; 15 avenue Laumière, F-75019 Paris (FR). **PAQUES, Frédéric** [FR/FR]; 92 Boulevard du Maréchal Joffre, F-92340 Bourg-la-Reine (FR). **PEREZ-MICHAUT, Christophe** [FR/FR]; 17 rue du Général Delestraint, F-75016 Paris (FR). **SMITH, Julianne** [US/FR]; 2 rue Mirbel, F-75005 Paris (FR). **SOURDIVE, David** [FR/FR]; 39 rue Chaptal, F-92300 Levallois-Perret (FR).

(74) Agent: **CABINET ORES**; 36 rue de St Pétersbourg, F-75008 Paris (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CUSTOM-MADE MEGANUCLEASE AND USE THEREOF

(57) Abstract: New rare-cutting endonucleases, also called custom-made meganucleases, which recognize and cleave a specific nucleotide sequence, derived polynucleotide sequences, recombinant vector cell, animal, or plant comprising said polynucleotide sequences, process for producing said rare-cutting endonucleases and any use thereof, more particularly, for genetic engineering, antiviral therapy and gene therapy.

WO 2004/067736 A2

CUSTOM-MADE MEGANUCLEASE AND USE THEREOF

The present invention relates to new rare-cutting endonucleases, also called custom-made meganucleases, which recognize and cleave a specific nucleotide sequence, to polynucleotide sequences encoding for said new rare-cutting endonucleases, to a vector comprising one of said polynucleotide sequences, to a cell, an animal, or a plant comprising one of said polynucleotide sequences or said rare-cutting endonucleases, to a process for producing one of said rare-cutting endonucleases and any use of the disclosed products and methods. More particularly, this invention contemplates any use such rare-cutting endonuclease for genetic engineering, antiviral therapy, genome therapy and gene therapy.

Homing endonucleases constitute a family of very rare-cutting endonucleases. It was first characterised at beginning of the Nineties by the use (*in vivo*) of the protein I-Sce I (Omega nuclease encoded by a mitochondrial group I intron of the yeast *Saccharomyces cerevisiae*). Homing endonucleases encoded by introns ORF, independent genes or intervening sequences (inteins) present striking structural and functional properties that distinguish them from "classical" restriction enzymes (generally from bacterial system R/MII). They have recognition sequences that span 12—40 bp of DNA, whereas "classical" restriction enzymes recognise much shorter stretches of DNA, in the 3—8 bp range (up to 12 bp for rare-cutter). Therefore, the homing endonucleases present a very low frequency of cleavage, even in the human genome.

Furthermore, general asymmetry of homing endonuclease target sequences contrasts with the characteristic dyad symmetry of most restriction enzyme recognition sites. Several homing endonucleases encoded by introns ORF or inteins have been shown to promote the homing of their respective genetic elements into allelic intronless or inteinless sites. By making a site-specific double-strand break in the intronless or inteinless alleles, these nucleases create recombinogenic ends, which engage in a gene conversion process that duplicates the coding sequence and leads to the insertion of an intron or an intervening sequence at the DNA level.

Homing endonucleases fall into 4 separated families on the basis of pretty well conserved amino acids motifs. For review, see Chevalier and Stoddard

(2001, Nucleic Acids Research, 29, 3757-3774). One of them is the dodecapeptide family (dodecamer, DOD, D1-D2, LAGLI-DADG, P1-P2). This is the largest family of proteins clustered by their most general conserved sequence motif: one or two copies (vast majority) of a twelve-residue sequence: the di-dodecapeptide. Homing endonucleases with one dodecapetide (D) are around 20 kDa in molecular mass and act as homodimer. Those with two copies (DD) range from 25 kDa (230 AA) to 50 kDa (HO, 545 AA) with 70 to 150 residues between each motif and act as monomer. Cleavage is inside the recognition site, leaving 4 nt staggered cut with 3'OH overhangs. I-Ceu I and I-Cre I illustrate the homing endonucleases with one Dodecapeptide (mono dodecapeptide). I-Dmo I, I-Sce I, PI-Pfu I and PI-Sce I illustrate homing endonucleases with two Dodecapeptide motifs. Structural models using X-ray crystallography have been generated for I-Cre I, I-Dmo I, PI-Sce I, PI-Pfu I. structures of I-Cre I bound to its DNA site have also been elucidated leading to a number of predictions about specific protein-DNA contacts. Seligman et al (Nucleic Acids Research, 2002, 30, 3870-3879) tests these predictions by analysing a set of endonuclease mutants and a complementary set of homing site mutants. In parallel, Gruen et al (Nucleic Acids Research, 2002, 30, e29) developed an *in vivo* selection system to identify DNA target site variants that are stik4 by wild-type homing endonucleases.

Endonucleases are requisite enzymes for today's advanced genetic engineering techniques, notably for cloning and analyzing genes. Homing endonucleases are very interesting as rare-cutter endonucleases because they have a very low recognition and cleavage frequency in large genome due to the size of their recognition site. Therefore, the homing endonucleases are used for molecular biology and for genetic engineering.

More particularly, homologous recombination provides a method for genetically modifying chromosomal DNA sequences in a precise way. In addition to the possibility of introducing small precise mutations in order to alter the activity of the chromosomal DNA sequences, such a methodology makes it possible to correct the genetic defects in genes which cause disease. Unfortunately, current methods for achieving homologous recombination are inherently inefficient, in that

homologous recombination-mediated gene repair can usually be achieved in only a small proportion of cells that have taken up the relevant "targeting or correcting' DNA. For example, in cultured mammalian cells, such recombinational events usually occur in only one in ten thousand transfected cells.

It has been shown that induction of double stranded DNA cleavage at a specific site in chromosomal DMA induces a cellular repair mechanism which leads to highly efficient recombinational events at that locus. Therefore, the introduction of the double strand break is accompanied by the introduction of a targeting segment of DNA homologous to the region surrounding the cleavage site, which results in the efficient introduction of the targeting sequences into the locus (either to repair a genetic lesion or to alter the chromosomal DNA in some specific way). Alternatively the induction of a double stranded break at a site of interest is employed to obtain correction of a genetic lesion via a gene conversion event in which the homologous chromosomal DNA sequences from an other copy of the gene donates sequences to the sequences where the double stranded break was induced. This latter strategy leads to the correction of genetic diseases either in which one copy of a defective gene causes the disease phenotype (such as occurs in the case of dominant mutations) or in which mutations occur in both alleles of the gene, but at different locations (as is the case of compound heterozygous mutations). (See WO 96/14408 ; WO 00/46386 ; US 5,830,729 ; Choulika et al., Mol Cell Biol, 1995, 15, 1965-73; Cohen-Tannoudji et al., Mol Cell Biol, 15 1998, 18, 1444-8; Donoho et al, Mol Cell Biol ; Rouet et al, Mol Cell Biol, 1994, 14, 5096-106 ; the disclosure of which is incorporated herein by reference).

Unfortunately, this method of genome engineering by induction of homologous recombination by a double stranded break is limited by the introduction of a recognition and cleavage site of a natural meganuclease at the position where the recombinational event is desired.

Up today, in a first approach for generating new endonuclease, some chimeric restriction enzymes have been prepared through hybrids between a zinc finger DNA binding domain and the non-specific DNA-cleavage domain from the natural restriction enzyme Fok I (Smith et al., 2000, Nucleic Acids Res, 28, 3361-9 ;

Smith et al., 1999, Nucleic Acids Res., 27, 274-281v; Kim et al, 1996, Proc Natl Acad Sci USA, 93, 1156-60 ; Kim & Chandrasegaran, 1994, Proc Natl Acad Sci USA, 91, 883-7 ; WO 95/109233 ; W0/9418313).

Another approach consisted of embedding DNA binding and catalytic activities within a single structural unit, such as type II restriction endonuclease. However, efforts to increase the length of recognition sequence or alter the specificity of these enzymes have resulted in the loss of catalytic activity or overall diminution of specificity due to the tight interdependence of enzyme structure, substrate recognition and catalysis (Lanio et al, 2000, Protein Eng., 13, 275-281).

Based on homing endonuclease, Chevalier et al. (2002, Molecular Cell, 10, 895-905) have generated an artificial highly specific endonuclease by fusing domains of homing endonucleases I-Dmo I and I-Cre I. The resulting enzyme binds a long chimeric DNA target site and cleaves it precisely at a rate equivalent to its natural parents.

However, this experiment leads to one endonuclease with a new specificity but it is not applicable to find an endonuclease that recognizes and cleaves any desired polynucleotide sequence.

Although these efforts, there is still a strong need of new rare-cutting endonucleases with new sequence specificity for the recognition and cleavage.

The present invention concerns a method for producing a custom-made meganuclease able to cleave a targeted DNA sequence derived from an initial meganuclease. This method comprises the steps of preparing a library of meganuclease variants and selecting the variants able to cleave the targeted DNA sequence.

In a first embodiment of the method for producing a custom-made meganuclease, the initial meganuclease is a natural meganuclease. Alternatively, said initial meganuclease is not a natural one. Preferably, said initial meganuclease is a homing endonuclease, more preferably a LAGLIDADG homing endonuclease. In a more preferred embodiment, said LAGLIDADG homing endonuclease is I-Cre I.

In a second embodiment of the method for producing a custom-made meganuclease, the library of meganuclease variants is generated by targeted mutagenesis, by random mutagenesis, by DNA shuffling, by directed mutation or by a combination thereof. Preferably, said library is generated by targeted mutagenesis. Said targeted mutagenesis is performed in meganuclease segments interacting with the DNA target, and more preferably introduced at the positions of the interacting amino acids. Optionally, the amino acids present at the variable positions comprise or are selected from the group consisting of D, E, H, K, N, Q, R, S, T, Y.

In a particular embodiment of the present invention, a library of I-Cre I variants is prepared by introducing amino acid diversity in positions selected from the group consisting of: Q26, K28, N30, S32, Y33, Q38, Q44, R68, R70 and T140. Preferably, a library of I-Cre I variants is prepared by introducing diversity in positions : a) Q26, K28, N30, Y33, Q38, Q44, R68, R70, T140 ; b) Q26, K28, N30, Y33, Q38, Q44, R68, R70 ; c) Q26, K28, N30, Y33, Q44, R68, R70 ; or d) Q26, K28, Y33, Q38, Q44, R68, R70. More preferably, a library of I-Cre I variants is prepared by introducing diversity in positions Q26, K28, N30, Y33, Q38, Q44, R68, and R70.

In a third embodiment of the method for producing a custom-made meganuclease, said selection of the variants able to cleave the targeted DNA sequence or a part thereof comprises the following steps:

a) a selection step for the binding ability, a screening step for the binding ability, a selection for the cleavage activity, and a screening step for the cleavage activity;

b) a selection step for the binding ability, a screening step for the binding ability, and a screening step for the cleavage activity;

c) a selection step for the binding ability, a selection for the cleavage activity, and a screening step for the cleavage activity; or,

d) a screening step for the binding apathy and a screening step for the cleavage.

Preferably, said selection of the variants able to cleave the targeted DNA sequence or a part thereof comprises a selection step for the binding ability, a

selection for the cleavage activity, and a screening step for the cleavage activity Optionally, a screening assay for the binding ability after a selection step based on the binding capacity can be done in order to estimate the enrichment of the library for meganuclease variants presenting a binding capacity.

Preferably, the selection and the screening based on the binding ability use the phage display technology. Preferably, the selection based on the cleavage activity uses a test in which the cleavage leads to either the activation of a positive selection marker or the inactivation of a negative selection marker. Preferably, the screening based on the cleavage activity uses a test in which the cleavage leads to a) the activation of a positive selection marker or a reporter gene; or b) the inactivation of a negative selection marker or a reporter gene.

Therefore, one object of the present invention is custom-made meganuclease produced by the above-mentioned method, a polynucleotide encoding said custom made meganuclease and any use thereof. Furthermore, the invention concerns a cell, an animal or a plant comprising said custom-made meganuclease or a polynucleotide encoding said custom-made meganuclease.

The invention concerns the use of a custom-made meganuclease for molecular biology, for *in vivo* or *in vitro* genetic engineering for *in vivo* or in vitro genome engineering for antiviral therapy, for genome therapy or for gene therapy.

More particularly, the invention concerns the use of a custom-made meganuclease for introducing a double-stranded break in a site of interest comprising the recognition and cleavage site of said meganuclease, thereby inducing a DNA recombination event, preferably a homologous recombination event, a DNA loss or cell death.

In a first embodiment of a method of genetic engineering, a custom-made meganuclease introduces a double-stranded break in a site of interest located on a vector and comprising the recognition and cleavage site of said meganuclease, thereby inducing a homologous recombination with another vector presenting homology with the sequence surrounding the cleavage site.

In a second embodiment of a method of genome engineering, the method comprises the following steps: 1) introducing a double-stranded break at the genomic locus comprising at least one recognition and cleavage site of a custom-made meganuclease according to the present invention; 2) providing a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus.

In a third embodiment of a method of genome engineering, the method comprises the following steps: 1) introducing a double-stranded break at the genomic locus comprising at least one recognition and cleavage site of a custom-made meganuclease according to the present invention; 2) maintaining under conditions appropriate for homologous recombination with chromosomal DNA sharing homologies to regions surrounding the cleavage site.

These methods of genetic and genome engineering could be used for repairing a specific sequence, modifying a specific sequence, for attenuating or activating an endogenous gene of interest, for introducing a mutation into a site of interest, for introducing an exogenous gene, for inactivating or deleting an endogenous gene or a part thereof, for translocating a chromosomal arm, or for killing the cell. The invention relates to the resulting cells and their uses.

Therefore, the invention concerns the use of at least one custom-made meganuclease according to the present invention to repair a specific sequence, to restore a functional gene in place of a mutated one, to modify a specific sequence, to attenuate or activate an endogenous gene of interest, to introduce a mutation into a site of interest, to introduce an exogenous gene or a part thereof, and to inactivate or delete an endogenous gene or part thereof, to translocate a chromosomal arm, or to leave the DNA unrepaired and degraded, by exposing cells, animals, or plants to said meganuclease. Optionally, said cells, animals, or plants are further exposed to a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus.

The invention also concerns a composition comprising at least one custom made meganuclease according to the present invention. Said composition is used for repairing a specific sequence, modifying a specific sequence for attenuating

or activating an endogenous gene of interest, for introducing a mutation into a site of interest, for introducing an exogenous gene or a part thereof, for inactivating or deleting an endogenous gene or a part thereof, to translocate a chromosomal arm, or to leave the DNA unrepaired and by exposing cells, animals, or plants to said meganuclease. Optionally, said composition can further comprise a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus.

The invention also relates to a method for treating or prophylaxis of a genetic disease in an individual in need thereof comprising (a) inducing in cells of the individual a double stranded cleavage at a site of interest comprising at least one recognition and cleavage site of a custom-made meganuclease according to the present invention, and introducing into the individual a targeting DNA, wherein said targeting DNA comprises (1) DNA sharing homologies to the region surrounding the cleavage site and (2) DNA which will be used to repair the site of interest in the chromosomal DNA.

In a second embodiment, the method for treating or prophylaxis of a genetic disease in an individual in need thereof comprises inducing in cells of the individual a double stranded break at a site of interest comprising at least one recognition and cleavage site of a custom-made meganuclease according to the present invention under conditions appropriate for the DNA homologous to the region surrounding the site of cleavage to be used in order to repair the site of interest.

The present invention further relates to the resulting cells and their uses, such as for treatment or prophylaxis of a disease or disorder in an individual.

The present invention concerns the use of at least one custom-made meganuclease according to the present invention to prevent, ameliorate or cure a genetic disease by exposing cells, animals or patients to said meganuclease. Optionally, said cells, animals or patients are also exposed to a targeting DNA construct comprising the sequence which repairs the site of interest flanked by sequences sharing homologies to the targeted locus.

The invention concerns a composition comprising at least one custom-made meganuclease according to the present invention. Preferably said

composition is used for preventing, ameliorating or curing a genetic disease by exposing cells, plants, animals or patients to said composition. Optionally, said composition can further comprise the targeting DNA construct comprising the sequence which repairs the site of interest flanked by sequences sharing homologies to the targeted locus.

The custom meganucleases according to the present invention can also be used as therapeutics in the treatment of diseases caused by infectious agents that present a DNA intermediate. It is therefore an object of the present invention to use at least one custom-made meganuclease according to the present invention to prevent, ameliorate or cure infection by an infectious agent by exposing said infectious agent and/or infected cells, animals, plants or patients to said meganuclease, the DNA target sequence of which being present in genome of said infectious agent. Preferably, said infectious agent is a virus.

Another object of the present invention is to use at least one meganuclease according to the present invention for inactivating or deleting an infectious agent in biologically derived products and products intended for biological uses by treating the products with said meganucleases. Preferably, said infectious agent is a virus.

A further object of the invention is a composition comprising at least one custom-made meganuclease according to the present invention for preventing, ameliorating or curing an infection by an infectious agent by exposing the infectious agent or the infected cells, animals or patients to said composition. Preferably, said infectious agent is a virus.

An additional object of the invention is to provide compositions comprising at least one meganuclease according to the present invention for inhibiting propagation of an infectious agent, inactivating or deleting an infectious agent in biologically derived products or products intended for biological use, or for disinfecting an object. Preferably, said infectious agent is a virus.

The invention also relates to a method for treating or prophylaxis of an infection by an infectious agent in an individual in need thereof comprising (a) introducing into individual's cells of the individual at least one custom-made

meganuclease presenting a recognition and cleavage site in the infectious agent sequence, and (b) inducing a double-strand break at said recognition and cleavage site, thereby leading to a recombination event resulting in inactivation or deletion of the infectious agent.

- Definitions

In the present application, by "meganuclease" is intended a double-stranded endonuclease having a polynucleotide recognition site of 14-40 bp. Said meganuclease is either monomeric or dimeric. Therefore, the meganuclease are also called rare-cutting or very rare cutting endonuclease. The homing endonucleases are one type of meganucleases.

By " custom-made meganuclease" is intended a meganuclease derived from an initial meganuclease presenting a recognition and cleavage site different from the site of the initial one. By " different" is intended that the custom-made meganuclease cleaves the site with an efficacity at least 10 fold more than the natural meganuclease, preferably at least 50 fold, more preferably at least 100 fold. The initial meganuclease can be a natural meganuclease or a modified one. By "natural"refers to the fact that an object can be found in nature. For example, a meganuclease that is present in an organism, that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is natural.

"Identity" refers to sequence identity between two nucleic acid molecules or polypeptides. Identity can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base, then the molecules are identical at that position. A degree of similarity or identity between nucleic acid or amino acid sequences is a function of the number of identical or matching nucleotides at positions shared by the nucleic acid sequences. Various alignment algorithms and/or programs may be used to calculate the identity between two sequences, including FASTA, or BLAST which are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.), and can be used with, e.g., default settings.

By " homologous " is intended a sequence with enough identity to another one to lead to a homologous recombination between sequences, more particularly having at least 95 % identity, preferably 97 %, and more preferably 99%.

The term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors A vector according to the present invention comprises, but is not limited to, a YAC (yeast artificial chromosome), a BAC (bacterial artificial), a baculovirus vector, a phage, a phagemid, a cosmid, a viral vector, a plasmid, a RNA vector or a linear or circular DNA or RNA molecule which may consist of chromosomal, non chromosomal, semi-synthetic or synthetic DNA. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. Large numbers of suitable vectors are known to those of skill in the art and commercially available, such as the following bacterial vectors: pQE7O, pQE6O, pQE-9 (Qiagen), pbs, pDIO, phagescript, psiXl74. pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); pWLNEO,pSV2CAT, pOG44, pXTI, pSG (Stratagene); pSVK3, pBPV, pMSG, pSVL (Pharmacia); pQE-30 (QlAexpress), pET (Novagen).

Viral vectors include retrovirus, adenovirus, parvovirus (e. g., adenoassociated viruses), coronavirus, negative strand RNA viruses such as orthomyxovirus (e. g., influenza virus), rhabdovirus (e. g., rabies and vesicular stomatitis virus), paramyxovirus (e. g. measles and Sendai), positive strand RNA viruses such as picornavirus and alphavirus, and double stranded DNA viruses including adenovirus, herpesvirus (e. g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomegalovirus), and poxvirus (e. g., vaccinia, fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, and hepatitis virus, for example. Examples of retroviruses include: avian leukosis-

sarcoma, mammalian C-type, B-type viruses, Dtype viruses, HTLV-BLV group, lentivirus, spumavirus (Coffin, J. M., Retroviridae: The viruses and their replication, In Fundamental Virology, Third Edition, B. N. Fields, et al., Eds., Lippincott-Raven Publishers, Philadelphia, 1996). Other examples include murine leukemia viruses, murine sarcoma viruses, mouse mammary tumor virus, bovine leukemia virus, feline leukemia virus, feline sarcoma virus, avian leukemia virus, human T-cell leukemia virus, baboon endogenous virus, Gibbon ape leukemia virus, Mason Pfizer monkey virus, simian immunodeficiency virus, simian sarcoma virus, Rous sarcoma virus and lentiviruses. Other examples of vectors are described, for example, in McVey et al., US 5,801,030, the teachings of which are incorporated herein by reference.

Vectors can comprise selectable markers (for example, neomycin phosphotransferase, histidinol dehydrogenase, dihydrofolate reductase, hygromycin phosphotransferase, herpes simplex virus thymidine kinase, adenosine deaminase, glutamine synthetase, and hypoxanthine-guanine phosphoribosyl transferase for eukaryotic cell culture ; TRP1 for *S. cerevisiae*; tetracycline, rifampicin or ampicillin resistance in *E. coli*; etc...). However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become and which become known in the art subsequently hereto.

The phrases "site of interest", "target site" and "specific site", as used herein, refer to a distinct DNA location, preferably a chromosomal location, at which a double stranded break (cleavage) is to be induced by the meganuclease.

As used herein, the term "individual" includes mammals, as well as other vertebrates (e.g., birds, fish and reptiles). The terms "mammal" and "mammalian", as used herein, refer to any vertebrate animal, including monotremes, marsupials and placental, that suckle their young and either give birth to living young (eutharian or placental mammals) or are egg-laying (metatharian or nonplacental mammals). Examples of mammalian species include humans and other primates (e.g., monkeys, chimpanzees), rodents (e.g., rats, mice, guinea pigs) and ruminants (e.g., cows, pigs, horses).

By "genetic disease" is intended any disease, partially or completely, directly or indirectly, due to an abnormality in one or several genes. Said abnormality can be a mutation, an insertion or a deletion. Said mutation can be a punctual muta-

tion. Said abnormality can affect the coding sequence of the gene or its regulatory sequence. Said abnormality can affect the structure of the genomic sequence or the structure or stability of the encoded mRNA. Said genetic disease can be recessive or dominant. Such genetic disease could be, but are not limited to, cystic fibrosis, Huntington's chorea, familial hyperchoiesterolemia (LDL receptor defect), hepatoblastoma, Wilson's disease, congenital hepatic porphyrias, inherited disorders of hepatic metabolism, Lesch Nyhan syndrome, sickle cell anemia, thalassaemias, xeroderma pigmentosum, Fanconi's anemia, retinitis pigmentosa, ataxia telangiectasia, Bloom's syndrome, retinoblastoma, Duchenne's muscular dystrophy, and Tay-Sachs disease.

- Generation of Meganuclease variants

The present invention concerns a method to produce a custom-made meganuclease specific to a targeted DNA sequence derived from an initial meganuclease by the introduction of diversity. Optionally, said initial meganuclease is a natural meganuclease. This method comprises the steps of preparing a library of meganuclease variants and isolating, by selection and/or screening, the variants able to bind andl/or cleave the targeted DNA sequence or a part thereof.

The diversity could be introduced in the meganuclease by any method available for the man skilled in the art. Preferably, the diversity is introduced by targeted mutagenesis (i.e. cassette mutagenesis, oligonucleotide directed codon mutagenesis, targeted random mutagenesis), by random mutagenesis (i.e. mutator strains, Neurospora crassa system (US 6,232,112; WO01/70946, error-prone PCR), by DNA shuffling, by directed mutation or a combination of these technologies (See Current Protocols in Molecular Biology, Chapter 8 "Mutagenesis in cloned DNA", Eds Ausubel et al., John Wiley and Sons). The meganuclease variants are preferably prepared by the targeted mutagenesis of the initial meganuclease. The diversity is introduced at positions of the residues contacting or interacting directly or indirectly with the DNA target. The diversity is preferably introduced in regions interacting with the DNA target, and more preferably introduced at the positions of the interacting amino acids. In libraries generated by targeted mutagenesis, the 20 amino acids can be introduced at the chosen variable positions. Preferably, the amino acids present at the

variable positions are the amino acids well-known to be generally involved in protein-DNA interaction. More particularly, these amino acids are generally the hydrophilic amino acids. More preferably, the amino acids present at the variable positions comprise D, E, H, K, N, Q, R, S, T, Y. Optionally, the amino acids present at the variable positions are selected from the group consisting of D, E, H, K, N, Q, R, S, T, Y. Synthetic or modified amino acids are also contemplated in the present invention.

One preferred way to generate a directed library is the use of degenerated codons at the positions where diversity has to be introduced. Several types of degenerated codons could be used. A degenerated codon N N K ([ATCG] [ATCG] [TG]) leads to 32 different codons encoding the 20 amino acids and one stop. A degenerated codon N V K ([ATCG] [ACG] [TG]) leads to 24 different codons encoding the 15 amino acids and one stop. A degenerated codon V V K ([ACG] [ACG] [TG]) leads to 18 different codons encoding the 12 amino acids (A, D, E, G, H, K, N, P, Q, R, S, T) and no stop. A degenerated codon R V K ([AG] [ACG] [TG]) leads to 12 different codons encoding the 9 amino acids (A, D, E, G, K, N, R, S, T). Preferably, a degenerated codon V V K ([ACG] [ACG] [TG]) leading to 18 different codons encoding the 12 amino acids (A, D, E, G, H, K, N, P, Q, R, S, T) is used for generating the library. Indeed, the V V K degenerated codon does not contain any stop codon and comprises all the hydrophilic amino acids.

If a directed library is generated, knowledge on amino acids interacting with the DNA target is useful. This knowledge could be provided, for example, by X-ray cristallography, Alanine scanning, or cross-linking experiments. The amino acids interacting with the DNA target can also be deduced by sequence alignment with a homologous protein.

The custom-made meganuclease is derived from any initial meganuclease. By initial meganuclease is intended a natural one or a modified one. Said modified one can be derived from natural ones by the hybrid generation or by a modification of physico-chemical properties of a natural one. Optionally, the initial meganuclease is selected so as its natural recognition and cleavage site is the closest to the targeted DNA site. Preferably, the initial meganuclease is a homing endonuclease, as specified, in the here above definitions. Homing endonucleases fall into 4 separated families on the basis of well conserved amino acids motifs, namely the LAGLIDADG

family, the GIY-YIG family, the His-Cys box family, and the HNH family (Chevalier et al., 2001, N.A.R, 29, 3757-3774).

The detailed three-dimensional structures of several homing endonucleases are known, namely I-Dmo I, PI-Sce I, PI-Pfu I, I-Cre I, I-Ppo I, and a hybrid homing endonuclease I-Dmo I / I-Cre I called E-Dre I (Chevalier et al., 2001, Nat Struct Biol, **8**, 312-316; Duan et al., 1997, Cell, **89**, 555-564; Heath et al., 1997, Nat Struct Biol, **4**, 468-476; Hu et al., 2000, J Biol Chem, **275**, 2705-2712; Ichiyanagi et al., 2000, J Mol Biol, **300**, 889-901; Jurica et al., 1998, Mol Cell, **2**, 469-476; Poland et al., 2000, J Biol Chem, **275**, 16408-16413; Silva et al., 1999, J Mol Biol, **286**, 1123-1136 ; Chevalier et al., 2002, Molecular Cell, 10, 895-905).

The LAGLIDADG family is the largest family of proteins clustered by their most general conserved sequence motif: one or two copies of a twelve-residue sequence: the di-dodecapeptide, also called LAGLIDADG motif. Homing endonucleases with one dodecapeptide (D) are around 20 kDa in molecular mass and act as homodimer. Those with two copies (DD) range from 25 kDa (230 AA) to 50 kDa (HO, 545 AA) with 70 to 150 residues between each motif and act as monomer. Cleavage is inside the recognition site, leaving 4 nt staggered cut with 3'OH overhangs. I-Ceu I, and I-Cre I illustrate the homodimeric homing endonucleases with one Dodecapeptide motif (mono-dodecapeptide). I-Dmo I, I-Sce I, PI-Pfu I and PI-Sce I illustrate monomeric homing endonucleases with two Dodecapeptide motifs.

The initial LAGLIDADG homing endonuclease can be selected from the group consisting of : I-Sce I, I-Chu I, I-Dmo I, I-Cre I, I-Csm I, PI-Sce I, PI-Tli I, PI-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, PI-Civ I, PI-Ctr I, PI-Aae I, PI-Bsu I, PI-Dha I, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu I, PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, PI-Tag I, PI-Thy I, PI-Tko I, and PI-Tsp I ; preferably, I-Sce I, I-Chu I, I-Dmo I, I-Cre I, I-Csm I, PI-Sce I, PI-Pfu I, PI-Tli I, PI-Mtu I, and I-Ceu I; more preferably, I-Dmo I, I-Cre I, PI-Sce I, and PI-Pfu I ; still more preferably I-Cre I.

The four structures of LAGLIDADG homing endonucleases, namely those of I-Dmo I, PI-Sce I, PI-Pfu I, and I-Cre I, reveal the functional significance of the LAGIDADG motif, and the nature of the DNA-binding interface. The core $\alpha\ \beta\ \beta$

α β β α fold of the homodimer homing endonuclease is repeated twice in the monomer homing endonuclease and confers upon the monomer a pseudo-dimeric structure. The first α-helix of each domain or subunit contains the defining LAGLIDADG motif. The two LAGLIDADG helices of each protein form a tightly packed dimer or domain interface. The DNA binding interface is formed by the four β-strands of each domain or subunit that fold into an antiparallel β-sheet. A minimal DNA binding moiety could be defined in the LAGLIDADG homing endonucleases as a β-hairpin (2 β-strands connected by a loop or turn), two such β-hairpins being connected into the 4-stranded β-sheet.

Each domain or subunit interacts with a half recognition site. The « external » quarter recognition site can be defined by its interaction with only one of the 2 β-hairpins of each domain or subunit.

Therefore, meganuclease variants derived from LAGLIDADG homing endonuclease can be fragmented in several directed libraries. This fragmented approach for the evolution of an initial meganuclease allows the introduction of a greater diversity (more amino acids at a position and/or more diversificated positions). In each library, the diversity is introduced only in the region involved in the interaction with a half or a quarter recognition site, the targeted DNA being modified only for the part interacting with the region comprising the introduced diversity. More particularly, if a new half site is searched for, then the diversity is preferably introduced in the 4-stranded β-sheet of one domain or subunit, more preferably at the positions of the DNA interacting amino acids in this structure. If a new quarter site is searched for, then the diversity is introduced in the corresponding β-hairpin, more preferably at the positions of the DNA interacting amino acids of this structure.

Preferably, a set of libraries covers the entire targeted DNA site. Hence, if the libraries comprise diversity only in the region interacting with a half-site, at least two libraries, preferably two, are necessary. However, if the initial meganuclease is a dimer, one library is enough with a half-site approach. If the libraries comprise diversity only in the region interacting with a quarter site, at least four libraries, preferably four, are necessary. If the initial meganuclease is a dimer, two libraries can be enough with a quarter site approach.

After the selection or screening of the primary libraries, the selected elements from the primary libraries are fused or combined in a subsequent library for a new cycle of selection. For example, two libraries can be fused by shuffling. A new cycle of selection could be then done on the whole targeted DNA site. Optionally, the new cycle of selection can be done on a half targeted DNA site if the first libraries are based on a quarter site. Subsequently, the results of the selection and/or screening of the half site are combined to give a final library which can be screened for the whole targeted DNA site.

Alternatively, the best elements from each libraries are joined together in order to obtain a meganuclease able to bind and cleave the targeted DNA site.

In an other approach, a library with diversity located only in the region involved in the interaction with a half or a quarter recognition site is prepared. Then, after selection or screening of this library, the selected elements from the library are modified such as to introduce diversity in another region involved in the interaction with recognition site, leading to a subsequent library. Libraries are generated until the complete targeted DNA site is bound and cleaved by the selected meganuclease.

More specifically, for the dimeric homing endonuclease (such as I-Cre I and I-Ceu I), a library can be generated by introducing diversity only in the region interacting with a half-site, a half site corresponding to one monomer of the initial homing endonuclease. This library can be used for selection and/or screening on each half sites of the target DNA sequence. When positive elements from the library have been selected for each half site, a variant for the first half site and a variant for the other half site are brought together for binding and cleaving the whole target DNA sequence. Alternatively, the positive variants can be introduced in a single chain meganuclease structure. As described in Example 1, a single chain meganuclease is an enzyme in which the two monomers of the initial dimeric homing endonuclease are covalently bound by a linker.

If an approach by a quarter site is chosen from an initial dimer homing endonuclease, at least two libraries are generated by introducing diversity only in the region involved in the interaction with each quarter recognition sites. After

the selection or screening of the primary libraries, the selected variants from the primary libraries are fused in a subsequent library for a new cycle of selection on the half site. Alternatively, the best elements from each libraries are joined together to obtain a monomer able to bind the half site. Otherwise, a library with diversity only in the region involved in the interaction with a quarter recognition site is prepared. Then, after selection or screening of this library, the selected elements from the library are modified such as to introduce diversity in the region involved in the interaction with the other quarter site, leading to a subsequent library. The selection and/or screening of this second library lead to the variants monomer able to bind the half site. When positive elements from the library have been selected for each half sites, a variant for the first half site and a variant for the other half site are brought together for binding and cleaving the target DNA sequence. Alternatively, the positive variants can be introduced in a single chain meganuclease structure.

In a preferred embodiment, the present invention concerns a method to prepare a custom-made meganuclease which recognizes and cleaves a desired polynucleotide target is derived from the directed evolution of the homing endonuclease I-Cre I. As the homing endonuclease is a homodimer, the approach in this case is based either on the half recognition site or on the quarter site.

The directed evolution is based on a library of I-Cre I variants. These I-Cre I variants present a diversity of amino acids at several positions predicted to interact with the polynucleotide target.

The X-ray structure of I-Cre endonuclease with its DNA target predicted that the following positions are involved: Q26, K28, N30, S32, Y33, Q38, Q44, R68, R70 and T140. Seligman et al (supra) showed that the positions S32 and T140 appear to be relatively unimportant for DNA recognition.

In one embodiment of the present invention, a set of I-Cre I variants is prepared by introducing amino acid diversity in positions selected from the group consisting of : Q26, K28, N30, S32, Y33, Q38, Q44, R68, R70 and T140. In a preferred embodiment, a set of I-Cre I variants is prepared by introducing diversity in positions : a) Q26, K28, N30, Y33, Q38, Q44, R68, R70, T140 ; b) Q26, K28, N30, Y33, Q38, Q44, R68, R70 ; c) Q26, K28, N30, Y33, Q44, R68, R70 ; or d) Q26, K28,

Y33, Q38, Q44, R68, R70. Preferably, a set of I-Cre I variants is prepared by introducing diversity in positions Q26, K28, N30, Y33, Q38, Q44, R68, and R70.

Optionally, the residue D75 of I-Cre I could be mutated in an uncharged amino acid such as N. Indeed, this amino acid has an interaction with 2 residues which are preferably modified in the library. As this charge is present in the core of the structure, it could be preferable to abolish this charge.

If the evolution approach of the homing endonuclease I-Cre I is based on the quarter recognition site, replacing the DNA binding residues presented by a β-hairpin (within the 4-stranded b-sheet) is a practical solution. As those residues are part of an element with limited length (i.e. less than 25 residue), they can be mutated together at once, for example by cassette replacement. Visual inspection of structure 1g9y (I-CreI with its target double-stranded DNA) indicates that the first β-hairpin is a unique or major contributor to the recognition of the last six bases of the target (i.e. either bases -12 to -7 or bases +7 to +12). Thus replacing the sequence from residue S22 to residue Q44, more preferably from residue I24 to residue T42, should be sufficient to specify new interaction specificity for the last six bases of the target site. More preferably, the residues interacting directly with DNA should be modified: I24, Q26, K28, N30, S32, Y33, Q38, S40 and T42. Alternatively (or in addition), the turn at the middle of the β-hairpin, which interacts with the very end of the 24bp-long DNA target, may be replaced by a short and flexible loop that would be tolerant to DNA bases substitution. For example, residues 30 to 36 could be replaced by 2, 3, 4, 5 or 6 glycine residues. This strategy is worth testing with all meganucleases presenting a comparable 3D structure. The second hairpin could be replaced similarly as a single unit (from residue Y66 to I77). However, while this hairpin interacts predominantly with the internal quarter site (bases -6 to -1 or +1 to +6), other residues (i.e. S22, Q44 and T46) separated from the hairpin may play a role in directing the specificity of interaction. Thus, a library could be created by replacing residues Y66, R68, R70, V73, D75 and I77. In parallel, S22, Q44 and T46 may either be left untouched, replaced by small polar amino acids (G, S or T; more preferably S or T), or randomized to contribute to the library. Mutants selected from separate library (the first wherein randomized residues are I24, Q26, K28, N30, S32, Y33, Q38, S40 and T42 and the second wherein randomized residues are Y66, R68, R70,

V73, D75 and I77) can be combined together by standard DNA shuffling methods based on recombination at homologous DNA regions (i.e. the DNA coding for the region between residue 43 and residue 65 is strictly conserved). However, if the second library includes mutations of residues S22, Q44 and T46, recombination becomes impractical, and more classical DNA/protein engineering is required.

If the evolution approach of the homing endonuclease I-Cre I is based on the quarter recognition site, a library of I-Cre I variants is prepared by introducing diversity in positions selected from the group consisting of : a) I24, Q26, K28, N30, S32, Y33, Q38, S40 and T42; or b) Y66, R68, R70, V73, D75, and I77. In the alternative b), the diversity could be also introduced in positions selected from the group consisting of: S22, Q44, and T46.

Alternatively, a custom-made meganuclease which recognizes and cleaves a desired polynucleotide target could be prepared by the directed evolution of single chain I-Cre I endonuclease. A set of single-chain I-Cre I variants is prepared by introducing amino acid diversity in positions selected from the group consisting of: Q26, K28, N30, S32, Y33, Q38, Q44, R68, R70, Q123, K125, N127, S129, Y130, Q135, Q141, R165, R167.

- Selection and Screening

Two properties of the meganuclease can be used for the steps of selection and/or screening, namely the capacity to bind the targeted DNA sequence and the ability to cleave it.

The meganuclease variants can be selected and screened, or only screened. The selection and/or screening can be done directly for the ability of the meganuclease to cleave the targeted DNA sequence. Alternatively, the selection and/or screening can be done for the binding capacity on the targeted DNA sequence, and then for ability of the meganuclease to cleave it. Preferably, the method to prepare a custom-made meganuclease comprises or consists of the following steps:

a) a selection step for the binding ability, a screening step for the binding ability, a selection for the cleavage activity, and a screening step for the cleavage activity;

b) a selection step for the binding ability, a screening step for the binding ability, and a screening step for the cleavage activity;

c) a selection step for the binding ability, a selection for the cleavage activity, and a screening step for the cleavage activity;

d) a screening step for the binding ability and a screening step for the cleavage activity;

e) a selection step for and a screening step for the cleavage activity; or,

f) a screening step for the cleavage activity.

More preferably, the method to prepare a custom-made meganuclease comprises or consists of the following steps: a selection step for the binding ability, a selection for the cleavage activity, and a screening step for the cleavage activity. A screening assay for the binding ability after a selection step based on the binding capacity can be done in order to estimate the enrichment of the library for meganuclease variants presenting a binding capacity.

The selection and screening assays are performed on the DNA region in which a double stranded cleavage has to be introduced or a fragment thereof. Preferably, the targeted sequences comprise at least 15 nucleotides, preferably 18 to 40, more preferably 18 to 30 nucleotides. In case of dimeric meganuclease, the targeted DNA polynucleotide can be reduced to at least 8 nucleotides for binding only. Preferably, the targeted DNA polynucleotide length is less than 10 kb, preferably less than 3 kb, more preferably less than 1 kb. For the DNA binding assay, the targeted DNA polynucleotide length is preferably less than 500 bp, more preferably less than 200 bp.

Any targeted sequence can be used to generate a custom-made meganuclease able to cleave it according. Optionally, the targeted sequence is chosen such as to present the most identity with the original recognition and cleavage site of the initial meganuclease.

Therefore, the DNA region in which a double stranded break has to be introduced is analyzed to choose at least 1, 2, 3 or 5 sequences of at least 15 nucleotides length, preferably 18 to 40 nucleotides, more preferably 18 to 30 nucleotides, having at least 25 % identity, preferably 50 % identity and more preferably 75 % identity with the original recognition and cleavage site of the initial meganuclease.

The targeted DNA sequence is adapted to the type of meganuclease variants library. If the library is based on a half site approach, the targeted DNA sequence used for the selection / screening comprises one half original site and one half site of the desired DNA sequence. If the library is based on a quarter site approach, the targeted DNA sequence used for the selection / screening comprises three quarters of the original site and one quarter site of the desired DNA sequence.

The meganuclease variants resulting from the selection and/or screening steps could optionally be an input for another cycle of diversity introduction.

The positive meganuclease variants selected by the selection and/or screening steps are validated by in vitro and/or ex vivo cleavage assay.

- Selection and/or Screening Based on Binding Property of Meganuclease

The selection and screening of meganuclease variants based on the binding capacity has to be made in conditions that are not compatible with the cleavage activity. For example, most of homing endonucleases need manganese or magnesium for their cleavage activity. Therefore, the binding assays on this type of homing endonuclease variants are done without manganese or magnesium, preferably replaced by calcium.

- Selection based on binding property of meganuclease

The binding selection assay is based on the enrichment of the meganuclease variants able to bind the targeted DNA polynucleotide. Therefore, the meganuclease variants encoded by the library are incubated with an immobilized targeted DNA polynucleotide so that meganuclease variants that bind to the immobilized targeted DNA polynucleotide can be differentially partitioned from those that do not present any binding capacity. The meganuclease variants which are bound to the immobilized targeted DNA polynucleotide are then recovered and amplified for a subsequent round of affinity enrichment and amplification. After several rounds of affinity enrichment and amplification, the library members that are thus selected can be isolated. Optionally, the nucleotide sequences encoding the selected meganuclease variants are determined, thereby identifying of the meganuclease variants able to bind the targeted DNA sequence.

The selection of meganuclease variants requires a system linking genotype and phenotype such as phage display (WO91/17271, WO91/18980, and WO91/19818 and WO93/08278 ; the disclosures of which are incorporated herein by reference), ribosome display (Hanes & Plückthun, PNAS, 1997, vol. 94, 4937-4942; He & Taussig, Nucl. Acids Res. (1997) vol. 25, p 5132-5143) and mRNA-protein fusion (WO00/47775 ; US 5,843,701 ; Tabuchi et al FEBS Letters 508 (2001) 309-312; the disclosures of which are incorporated herein by reference).

Phage display involves the presentation of a meganuclease variant on the surface of a filamentous bacteriophage, typically as a fusion with a bacteriophage coat protein. The library of meganuclease variants is introduced into a phage chromosome or phagemid so as to obtain a protein fusion with a bacteriophage coat protein, preferably with the pIII protein. If the initial meganuclease is a homodimer, the monomer variants of the meganuclease are introduced so as to be displayed and the constant monomer can be introduced so as to be produced in the periplasm. The bacteriophage library can be incubated with an immobilized targeted DNA sequence so that elements able to bind the DNA are selected.

mRNA-protein fusion system opens the possibility to select among 10^{13} different meganuclease variants. This system consists in the creation of a link between the mRNA and the encoded protein via a puromycin at the 3' end of the mRNA which leads to a covalent mRNA-protein fusion at the end of the translation. Hence, a double-stranded DNA library comprising the coding sequence for the meganuclease variants is used regenerate mRNA templates for translation that contain 3' puromycin. The mRNA-puromycin conjugates are translated in vitro to generate the mRNA-meganuclease fusions. After cDNA synthesis, the fusions are tested for the ability to bind the immobilized targeted DNA polynucleotide. A PCR is then used to generate double-stranded DNA enriched in meganuclease variants presenting the binding capacity. If the initial meganuclease is a homodimer, the constant monomer can be introduced either as DNA or mRNA encoding this monomer or as a monomer protein. In this case, an approach with the single chain meganuclease will be preferably used.

Ribosome display involves a double-stranded DNA library comprising the coding sequence for the meganuclease variants that is used to generate

mRNA templates for translation. After a brief incubation, translation is halted by addition of Mg^{2+} and incubation al low temperature or addition of translation inhibitor. The ribosome complexes are then tested for the ability to bind immobilized targeted DNA polynucleotide. The selected mRNA is used to construct cDNA and a PCR generates double-stranded DNA enriched in meganuclease variants presenting the binding capacity. If the initial meganuclease is a homodimer, the constant monomer is introduced either as DNA or mRNA encoding this monomer or as a monomer protein. In this case, an approach with the single chain meganuclease will be preferably used.

The targeted DNA sequence can be immobilized on a solid support. Said solid support could be a column, paramagnetic beads or a well of a microplate. For example, the polynucleotides comprising the targeted DNA sequence present a ligand (such as a biotin) at one end, said ligand allowing the immobilization on a solid support bearing the target of the ligand (for example, streptavidin if biotin is used).

The selection of the meganuclease variants may usually be monitored by a screening assay based on the binding or cleavage capacity of these meganucleases. However, the selected meganuclease variants can be also directly introduced in a selection step based on the cleavage capacity.

- Screening based on binding property of meganuclease

In order to perform the screening assay, the selected meganuclease variants need to be cloned. If the selection was done with the phage display system, the clone encoding each meganuclease variants can be easily isolated. If the selection was done by mRNA-protein fusion or ribosome display, the selected meganuclease variants have to be subcloned in expression vector.

The screening assays are preferably performed in microplates (96, 384 or 1536 wells) in which the targeted DNA polynucleotides are immobilized. After expression of the meganuclease variants, these variants are incubated with the immobilized targeted DNA polynucleotides. The meganuclease variants expression can be performed either *in vivo* or *in vitro*, preferably by *in vitro* expression system. Preferably, the meganuclease variants are purified prior to the incubation with the targeted polynucleotide. The retained meganuclease variants are then detected. The detection could be done by several means well known by the man skilled in the art. For

example, if phages are used, the detection can be done with antibodies against phages (ELISA). Otherwise, the expression could be done in presence of S35 amino acids in order to obtain radioactive meganucleases. Thus, the binding is estimated by a radioactivity measurement. The invention also considers the others means of detection of DNA binding by meganuclease available to the man skilled in the art.

Optionally, the nucleotide sequences encoding the positively screened meganuclease variants are determined, thereby identifying of the meganuclease variants able to bind the targeted DNA sequence.

The positively screened meganuclease variants have to be tested for their cleavage capacity. Therefore, said meganuclease variants are incorporated in a cleavage selection and/or screening experiment, preferably an *in vivo* cleavage screening assay. Optionally, said meganuclease variants can be tested by an *in vitro* cleavage assay.

The screening assay can also be used only for estimate the enrichment in meganuclease variants presenting the binding capacity. This estimation helps to decide if a new round of selection based on the binding capacity is necessary or if the selected library can be submitted to a cleavage selection and/or screening, preferably an in vivo cleavage selection and/or screening.

- Selection and/or Screening Based on Cleavage Property of Meganuclease

The selection and screening of meganuclease variants based on the cleavage capacity has to be made in conditions compatible with the cleavage activity. The meganuclease variants used in the selection and/or screening based on cleavage capacity may be either the initial library of meganuclease variants or the meganuclease variants selected and/or screened for the binding activity.

If necessary, the selected and/or screened meganuclease variants are subcloned in an appropriate expression vector for the *in vitro* and *in vivo* cleavage assay. Such subcloning step can be performed in batch or individually. More particularly, if the initial meganuclease is a dimer, the subcloning step allows the introduction of the selected library(ies) in a single chain meganuclease structure. If two libraries have been selected and/or screened for two half recognition and cleavage

sites, the subcloning step allows to bring together the two selected libraries in a single chain meganuclease structure.

- Selection based on cleavage property of meganuclease

The general principle of an *in vivo* selection of the meganuclease variants based on their cleavage capacity is that the double-strand break leads to the activation of a positive selection marker or the inactivation of a negative selection marker.

If the selection is based on the inactivation of a negative selection marker, the method involves the use of cell containing an expression vector comprising the coding sequence for a negative selection marker and the targeted DNA sequence for the desired meganuclease and an expression vector comprising the library of meganuclease variants. Preferably said expression vector is a plasmid. Preferably said targeted DNA sequence is located either near the negative selection gene or in the negative selection gene, preferably between the promoter driving the expression of the negative selection and the ORF. The expression of the negative selection marker has to be conditional in order to keep the cell alive until the meganuclease variants have the opportunity to cleave. Such a conditional expression can be easily done with a conditional promoter. However, there are other conditional systems that could be used. The meganuclease variants are introduced in an expression cassette. The meganuclease encoding sequence can be operably linked to an inducible promoter or to a constitutive promoter. Of course, the promoter is compatible with the cell used in the assay. If the meganuclease variant has the capacity to cleave the targeted DNA, then the negative selection marker is inactivated, either by deleting the whole negative marker gene or a part thereof (coding sequence or promoter) or by degrading the vector. A culture in a negative selection condition allows the selection of the cell containing the meganuclease variants able to cleave the targeted DNA sequence.

The vector comprising the negative selection marker is preferably transfected before the introduction of the vector encoding the meganuclease variants. Optionally, the vector comprising the negative selection marker can be conserved in the cell in an episomal form. Alternatively, the vector comprising the negative selection marker and the vector encoding the meganuclease variants can be cotransfected

into the cell. The cell can be prokaryotic or eukaryotic. Preferably, the prokaryotic cell is *E. coli*. Preferably, the eukaryotic cell is a yeast cell. The negative selection marker is a protein directly or indirectly toxic for the cell. For example, the negative selection marker can be selected from the group consisting of toxins, translation inhibitors, barnase, and antibiotic for bacteria, URA3 with 5FOA (5-fluoro-orotic acid) medium and LYS2 with a α-AA medium (alpha-adipic acid) for yeast, and thymidine kinase for superior eukaryotic cells. For an example of negative marker selection, see Gruen et al., 2002, Nucleic Acids Research, 30, e29; the disclosure of which is incorporated herein by reference.

If the selection is based on the activation of a positive selection marker, the method involves the use of cell containing an expression vector comprising an inactive positive selection marker and the targeted DNA sequence for the desired meganuclease and an expression vector comprising the library of meganuclease variants. Optionally, the inactive positive selection marker, the targeted DNA sequence and the library of meganuclease variants can be on the same vector (See WO 02/44409). Preferably said expression vector is a plasmid. The meganuclease variants are introduced in an expression cassette. The meganuclease encoding sequence can be operably linked to an inducible promoter or to a constitutive promoter. Of course, the promoter is compatible with the cell used in the assay. For example, the positive selection marker can be an antibiotic resistance (e.g. tetracycline, rifampicin and ampicillin resistance) or an auxotrophy marker for bacteria, TRP1, URA3, or an auxotrophy marker for yeast, and neomycine et puromycine for superior eukaryotic cell. Optionally, the positive selection marker can be an auxotrophy marker compatible with both bacteria and yeast (e.g. URA3, LYS2, TRP1, and LEU2). The inactive positive selection marker gene and the targeted DNA sequence have to be arranged so that the double-strand break leads to a rearrangement of the marker in an active positive marker. Two kinds of repair processes can lead to an active positive selection marker, namely single-strand annealing (SSA) or gene conversion (GC).

The *in vivo* Single-strand annealing recombination test (SSA) is known by the man skilled in the art and disclosed for example in Rudin et al. (Genetics 1989, **122**, 519-534 ; Fishman-Lobell & Haber (Science 1992, **258**, 480-4); Lin et al (Mol. Cell. Biol., 1984, **4**, 1020-1034) and Rouet et al (Proc. Natl. Acad. Sci.

USA, 1994, **91**, 6064-6068); the disclosure of which are incorporated herein by reference.

To test the meganuclease variants, an *in vivo* assay based on SSA in a cell, preferably a bacterial or yeast cell has been developed. For instance, the method uses a yeast cell. This organism has the advantage that it recombines naturally its DNA via homologous recombination with a high frequency.

This *in vivo* test is based on the reparation by SSA of a positive selection marker induced by double-strand break generated by an active meganuclease variant. The target consists of a modified positive selection gene with an internal duplication separated by a intervening sequence comprising the targeted DNA sequence. The internal duplication should contain at least 50 bp, preferably at least 200 bp. The efficiency of the SSA test will be increased by the size of the internal duplication. The intervening sequences are at least the targeted DNA sequence. The intervening sequence can optionally comprise a selection marker, this marker allowing checking that the cell has not repaired the positive selection marker by a spontaneous recombination event. The positive selection marker gene is preferably operably linked to a constitutive promoter relating to the cell used in the assay. According to said assay method, the cell will be selected only if a SSA event occurs following the double-strand break introduced by an active meganuclease variant.

Optionally, each vector can comprise a selectable marker to ensure the presence of the plasmid in the cell. The presence of this selectable marker is preferable for the assay performed in yeast cell. For example, for yeast, a first construct comprising the target gene can comprise a Leu2 selectable marker allowing transformed yeast to grow on a synthetic medium that does not contain any Leucine and a second construct can comprise the Trp1 selectable marker allowing transformed yeast to grow on a synthetic medium that does not contain any tryptophane.

The vector comprising the positive selection marker is preferably transfected before the introduction of the vector encoding the meganuclease variants. Optionally, the vector comprising the positive selection marker can be conserved in the cell in an episomal form. Alternatively, the vector comprising the positive selection marker and the vector encoding the meganuclease variants can be cotransfected into the cell.

The *in vivo* selection of the meganuclease variants can also be performed with a gene conversion assay. For example, the selection vector comprises a first modified positive selection gene with a deletion or a mutation and an insertion of the targeted DNA sequence for the meganuclease at the place of the deletion. The positive selection gene can also be inactivated by the interruption of the gene by an insert comprising the targeted DNA sequence. The selection construct further comprises the segment of the positive selection marker gene which has been deleted flanked at each side by the positive selection marker gene sequences bordering the deletion. The bordering sequences comprise at least 100 bp of homology with the positive selection marker gene at each side, preferably at least 300 pb. The double-stand break generated by an active meganuclease variant in the targeted DNA sequence triggers on a gene conversion event resulting in a functional positive selection marker gene. This kind of assay is documented in the following articles: Rudin et al (Genetics 1989, **122**, 519-534), Fishman-Lobell & Haber (Science 1992, **258**, 480-4), Paques & Haber (Mol. Cell. Biol., 1997, **17**, 6765-6771), the disclosures of which are incorporated herein by reference.

Otherwise, the *in vivo* selection of the meganuclease variants can be performed through a recombination assay on chromosomic target. The recombination can be based on SSA or gene conversion mechanisms. The *in vivo* selection can be based on several SSA targets, preferably at least two SSA targets.

A first example based on SSA is the following. A modified positive selection gene with an internal duplication separated by an intervening sequence comprising the targeted DNA sequence for the desired meganuclease variant is introduced into the chromosome of the cell. The internal duplication should contain at least 50 bp, preferably at least 200 bp. The efficiency of the SSA test will be increased by the size of the internal duplication. The intervening sequence is at least the targeted DNA sequence. By transfecting the cell with an expression construct allowing the production of a meganuclease variant in the cell, the repair by homologous recombination of the double-strand break generated by an active meganuclease variant will lead to a functional positive selection marker gene.

Another example based on gene conversion is the following. A mutated non-functional positive selection marker gene comprising the targeted DNA

sequence for the desired meganuclease variant is introduced into the chromosome of the cell. Said targeted DNA sequence has to be in the vicinity of the mutation, preferably at less than 1 kb from the mutation, more preferably at less than 500 bp, 200 bp, or 100 pb surrounding the mutation. By transfecting the cell with a fragment of the functional positive selection marker gene corresponding to the mutation area and an expression construct allowing the production of a meganuclease variant in the cell, the repair by homologous recombination of the double-strand break generated by an active meganuclease variant will lead to a functional positive selection marker gene. Alternatively, the fragment of the functional positive selection marker allowing the repair can be integrated on the chromosome. This kind of assay is documented in the following articles : Rouet et al (Mol. Cell. Biol., 1994, **14**, 8096-8106) ; Choulika et al (Mol. Cell. Biol., 1995, **15**, 1968-1973); Donoho et al (Mol. Cell. Biol., 1998, **18**, 4070-4078); the disclosures of which are incorporated herein by reference.

The selected clones comprise a meganuclease variant presenting the capacity to cleave the targeted DNA sequence. It is preferable to validate the selection by a screening assay. This screening assay can be performed *in vivo* or *in vitro*, preferably *in vivo*.

Optionally, the nucleotide sequences encoding the positively screened meganuclease variants are determined, thereby identifying the meganuclease variants able to cleave the targeted DNA sequence.

- Screening based on cleavage property of meganuclease

In order to perform the screening assay, the selected meganuclease variants need to be cloned and the cleavage assay need to be performed individually for each clone.

The *in vivo* cleavage assay for the screening is similar to those used for the selection step. It can be based on the inactivation of either a negative selection marker or a reporter gene, or on the activation of either a positive selection marker or a reporter gene.

By reporter gene is intended any nucleic acid encoding a product easily assayed, for example β-galactosidase, luciferase, alkaline phosphatase, green fluorescent protein, tyrosinase, DsRed proteins. The reporter gene is preferably oper-

ably linked to a constitutive promoter relating to the cell used in the assay (for example CMV promoter).

Cells used for this screening assay can be prokaryotic, preferably *E. coli*, or eukaryotic, preferably a yeast cell or a mammalian cell. More particularly, it could be interesting to use mammalian cells for a validation of a positive meganuclease variant by an *ex vivo* cleavage assay

- In Vitro Cleavage Assay

The recognition and cleavage of the targeted DNA sequence or a part thereof by the meganuclease variants can be assayed by any method known by the man skilled in the art.

One way to test the activity of the meganuclease variants is to use an in vitro cleavage assay on a polynucleotide substrate comprising the targeted DNA sequence or a part thereof. Said polynucleotide substrate could be a synthetic target site corresponding to:

- the whole targeted DNA site ;
- a half targeted DNA site and a half original site ; or,
- a quarter targeted DNA site and three quarters original site.

Said polynucleotide substrate can be linear or circular and comprises preferably only one cleavage site. The assayed meganuclease variant is incubated with the polynucleotide substrate in appropriate conditions. The resulting polynucleotides are analyzed by any known method, for example by electrophoresis on agarose or by chromatography. If the polynucleotide substrate is a linearized plasmid, the meganuclease activity is detected by the apparition of two bands (products) and the disappearance of the initial full-length substrate band. Preferably, said assayed meganuclease variants are digested by proteinase K, for example, before the analysis of the resulting polynucleotides. For instance, the polynucleotide substrate is prepared by the introduction of a polynucleotide comprising the sequence of the target site in a plasmid by TA or restriction enzyme cloning, optionally followed by the linearization of the plasmid. Preferably, such linearization is not done in the surrounding of the targeted DNA sequence. See Wang et al, 1997, Nucleic Acid Research, 25, 3767-3776; See Examples, Materials & Methods "*in vitro* activity assays" section) and the characterization papers of the initial homing endonucleases.

Alternatively, such in vitro cleavage assay can be performed with polynucleotide substrates linked to fluorophores, such substrates comprising the targeted DNA sequence. These polynucleotide substrates are immobilized on a solid support. Said solid support is preferably a microplate (96, 384 or 1536 wells). For example, the polynucleotides comprising the targeted DNA sequence present a ligand (such as a biotin) at one end, said ligand allowing the immobilization on a solid support bearing the target of the ligand (for example, streptavidin if biotin is used). The end opposite to the immobilized end is linked to a fluorophore. Cleavage leads to loss of fluorescence by release of the fluorochrome from the solid support.

Otherwise, some *in vitro* cleavage assays can be based on the fluorescence quenching. A fluorophore (for example, FAM or TAMRA) and a quencher (for example, DABCYL) are located on the polynucleotide substrate such as the quencher inhibits the fluorescence emission. The quenching is abolished when the cleavage by the meganuclease variants occurs on the polynucleotide substrates. Several examples of this quenching assays are detailed in Eisenschmidt et al (2002, Journal of Biotechnology, 96, 185-191) and WO 02/42497, the disclosure of these documents are incorporated herein by reference.

- Custom-Made Meganucleases, Polynucleotides Encoding A Custom-Made Meganuclease, Vectors, Cells And Animals/Plants

The present invention concerns any custom-made meganuclease prepared by the method according to the present invention and any use of it. Optionally, said meganuclease comprises a purification tag.

The present invention concerns a recombinant polynucleotide encoding a custom-made meganuclease prepared by a method according to the present invention. The present invention concerns:

a) any vector comprising a polynucleotide sequence encoding a custom-made meganuclease according to the present invention;

b) any prokaryotic or eukaryotic cell comprising either a polynucleotide sequence encoding a custom-made meganuclease according to the present invention or a vector according a); and,

c) any non-human animal or plant comprising a polynucleotide sequence encoding a custom-made meganuclease according to the present invention, or a vector according a), or a cell according b).

As used herein, a cell refers to a prokaryotic cell, such as a bacterial cell, or eukaryotic cell, such as an animal, plant or yeast cell. A cell which is of animal or plant origin can be a stem cell or somatic cell. Suitable animal cells can be of, for example, mammalian, avian or invertebrate origin. Examples of mammalian cells include human (such as HeLa cells), bovine, ovine, caprine, porcine, murine (such as embryonic stem cells), rabbit and monkey (such as COS1 cells) cells. The cell may be an embryonic cell, bone marrow stem cell or other progenitor cell. Where the cell is a somatic cell, the cell can be, for example, an epithelial cell, fibroblast, smooth muscle cell, blood cell (including a hematopoietic cell, red blood cell, T-cell, B-cell, etc.), tumor cell, cardiac muscle cell, macrophage, hepatic cell, dendritic cell, neuronal cell (e.g. a glial cell or astrocyte), or pathogen-infected cell (e.g., those infected by bacteria, viruses, virusoids, parasites, or prions).

The cells can be obtained commercially or from a depository or obtained directly from an individual, such as by biopsy. The cell can be obtained from an individual to whom they will be returned or from individual of the same or different species. For example, nonhuman cells, such as pig cells, can be modified to include a DNA construct and then put into a human. Alternatively, the cell need to be isolated from the individual for example, it is desirable to deliver the vector to the individual in gene therapy.

The vector comprising a polynucleotide encoding a custom-made meganuclease contains all or part of the coding sequence for said meganuclease operably linked to one or more expression control sequences whereby the coding sequence is under the control of transcriptional signals to permit production or synthesis of said meganuclease. Therefore, said polynucleotide encoding a custom made meganuclease is comprised in an expression cassette. More particularly, the vector comprises a replication origin, a promoter operatively linked to said encoding polynucleotide, a ribosome site, an RNA-splicing site (when genomic DNA is used), a polyadenylation site and a transcription termination site. It also can comprise an

enhancer. Selection of the promoter will depend upon the desired route for expressing the meganuclease.

The invention concerns a method for producing a custom-made meganuclease comprising introducing an expression vector into a cell compatible with the element of said expression vector.

The polynucleotide sequence encoding the custom meganuclease can be prepared by any method known by the man skilled in the art.

- Use of the meganuclease according to the invention

The custom-made meganucleases according to the present invention are of great utility. Of course, these custom-made meganucteases are precious for molecular biology and for genetic engineering, antiviral therapy, genome therapy and gene therapy, more particularly according to the methods described in WO 96/14408, US 5,830,729, WO 00/46385, WO 00/46386, the disclosure of these documents being incorporated by reference.

The custom-made meganucleases with new specificity according to the present abolish the limiting step of introducing the recognition and cleavage site for a natural meganuclease in the method of genetic engineering involving meganucleases.

- Genetic Engineering and Gene Therapy

The custom-made meganuclease according to the present invention can be used in genetic engineering for the preparation of vector. *In vitro*, said meganuclease are useful when a rare-cutting endonuclease is necessary in the vector construction. Said custom-made meganuclease can also be used for *in vivo* vector construction. For example, if the recognition and cleavage site for a custom-made meganuclease is located on a vector, said meganuclease can be used to induce a homologous recombination with an other vector presenting homology with the sequence surrounding the cleavage site. Said vector can be a plasmid or a viral vector. Similarly, said custom made meganuclease can be used to delete a helper vector (generally a plasmid) in a transcomplementing cell line for the production of retroviruses, AAV or adenoviruses.

Genome engineering is the set of methods used to induce a change in the genetic program of a living cell and/or organism. The meganucleases obtained by the method of the present invention allows rational site directed modifications of cell genomes. The purpose of these techniques is to rewrite chromosomes precisely where they should be modified leaving the rest of the genome intact. Fields of applications of the genome engineering are multiple: animal models generation (knock-in or knock-out), protein production (engineering of production strains, protein production in plant and animals for protein production in milks), agricultural biotechnology (addition or removal of a trait, marker excision), modification and study of metabolic pathway, or therapy of genetic diseases or viral diseases.

A custom-made meganuclease according to the present invention can be used in a method of genome engineering comprising: 1) introducing a double-strand break at the genomic locus comprising at least one recognition and cleavage site of said meganuclease; and, 2) providing a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus. Indeed, shared DNA homologies are located in regions flanking upstream and downstream the site of the break in the targeting DNA construct and the DNA that might be introduced should be located between the two arms. Said meganuclease can be provided directly to the cell or through an expression vector comprising the polynucleotide sequence encoding said meganuclease and suitable for its expression in the used cell. Alternatively, the method of genome engineering comprises: 1) introducing a double-strand break at the genomic locus comprising at least one recognition and cleavage site of said meganuclease; 2) maintaining under conditions appropriate for homologous recombination with the chromosomal DNA homologous to the region surrounding the cleavage site. Any of these methods of genetic engineering could be used for repairing a specific sequence, modifying a specific sequence, for attenuating or activating an endogen gene of in for introducing a mutation into a site of interest, for introducing an exogenous gene or a part thereof, for inactivating or deleting an endogenous gene or a part thereof, for methylated or demethylating the CpG dinucleotides of a gene. The invention relates to the resulting cells and their uses.

The custom-made meganuclease according to the present invention could also be used for addition or substitution of telomer, for killing cells, for chromosomic translocation, for changing the chromatinization, or for chromosomic loss.

It is an object of the invention to use at least one custom-made meganuclease according to the present invention to repair a specific sequence, to modify a specific sequence, to attenuate or activate an endogenous gene of interest, to introduce a mutation into a site co interest, to introduce an exogenous gene or a part thereof, and to inactivate or delete an endogenous gene or a part thereof by exposing cells animals, plants to said meganuclease.

Another object of the invention is a composition comprising at least one custom-made meganuclease according to the present invention. Preferabiy said composition is used for repairing a specific sequence, modifying a specific sequence, for attenuating or activating an endogenous gene of interest, for introducing a mutation into a site of interest, for introducing an exogenous gene or a part thereof, for inactivating or deleting an endogenous gene or a part thereof by exposing cells, animals, or plants to said meganuclease. Preferably said composition comprises one custom-made meganuclease or two different custom-made meganucleases. Optionally, said composition can further comprise a targeting DNA construct comprising the sequence to be introduced flanked by homologous sequence to the targeted locus.

More particularly, the invention also relates to the use of a custom-made meganuclease obtained by the method according to the present invention in a method for treating or prophylaxis of a genetic disease in an individual in need thereof comprising (a) inducing in cells of the individual a double stranded cleavage at a site of interest comprising at least one recognitron and cleavage site of said meganuclease, and (b) introducing into the individual a targeting DNA, wherein said targeting DNA comprises (1) DNA sharing homologies to the region surrounding the cleavage site and (2) DNA which repairs the site of interest upon recombination between the targeting DNA and the chromosomal DNA. The targeting DNA is introduced into the individual under conditions appropriate for introduction of the targeting DNA into the site of interest. In a second embodiment the method for treating or prophylaxis of a

genetic disease in an individual in need thereof comprises inducing in cells of the individual a double stranded break at a site of interest comprising at least one recognition and cleavage site of said meganuclease under conditions appropriate for chromosomal DNA homologous to the region surrounding to be introduced or deleted into the site of interest and repair of the site of interest. Alternatively, cells can be removed from an individual to be treated, modified by the present method and reintroduced by autographt into the individual.

The invention relates to custom-made meganuclease obtained by the method according to the present invention in a method for correcting a genetic lesion or abnormality in chromosomal DNA of a cell comprising inducing in the cell double stranded break at a site of interest in the genetic lesion or abnormality comprising at least one recognition and cleavage site of said meganuclease under conditions appropriate for chromosomal DNA homologous to the region surrounding the site of cleavage to be introduced into the site of interest and correct the genetic lesion or abnormality. Here, too, the method can be carried out in cells present in an individual or in cells removed from the individual, modified and then returned to the individual (ex vivo).

The present invention further relates to the resulting cells and their uses, such as for treatment or prophylaxis of a condition or disorder in an individual (e.g., a human or other mammal or vertebrate). For example, cells can be produced (e.g., ex vivo) by the method described herein and then introduced into an individual using known methods. Alternatively, cells can be modified in the individual (without being removed from the individual).

It is therefore an object of the present invention to use at least one custom-made meganuclease according to the present invention to prevent, ameliorate or cure a genetic disease by exposing cells, animals or patients to said meganuclease. Preferably, the invention relates to the use of one custom-made meganuclease or two different custom-made meganucleases.

Another object of the invention is a composition comprising at least one custom-made meganuclease according to the present invention. Preferably said composition is used for preventing, ameliorating or cure a genetic disease by

exposing cells, animals or patients to said composition. Preferably said composition comprises at least one custom-made meganuclease. Optionally, said composition can further comprise a targeting DNA construct comprising the sequence to be introduced flanked by sequences homologous to the targeted locus.

Targeting DNA and/or custom-made meganucleases introduced into a cell, an animal, a plant or an individual as described above can be inserted in a vector..Vectors comprising targeting DNA and/or nucleic acid encoding a meganuclease can be introduced into a cell by a variety of methods (e.g., injection, transformation, transfection, direct uptake, projectile bombardment, liposomes). Meganucieases can be stably or transiently expressed into cells using expression vectors. Techniques of expression in eukaryotic cells are well known to those in the art. (See Current Protocols in Human Genetics: Chapter 12 "Vecto Therapy" & Chapter 13 "Delivery Systems for Gene Therapy"). Optionally, it may be preferable to incorporate a nuclear localization signal into the recombinant protein to be sure that it is expressed within the nucleus. Custom-made meganuclease can also be introduced into a cell according to methods generally known in the art which are appropriate for the particular meganuclease and cell type. Custom-made meganucleases according to the present invention can be introduced into cells using liposomes or by fusion to the membrane translocating peptide (Bonetta, 2002, The Sientist, 16, 38; Ford et al, Gene Ther., 2001,8, 1-4; Wadia & Dowdy, 2002, CurrOpin Biotechnol, 13, 52-56).

Once in the cell, the custom-made meganuclease and the vector comprising targeting DNA and/or nucleic acid encoding a custom-made meganuclease are imported or translocated by the cell from the cytoplasm to the site of action in the nucleus.

Custom-made meganucleases and vectors which comprise targeting DMA homologous to the region surrounding the cleavagé site and/or nucleic acid encoding a custom-made meganuclease can be introduced into an individual using routes of administration generally known in the art. Administration may be topical or internal, or by any other suitable avenue for introducing a therapeutic agent to a patient. Topical administration may be by application to the skin, or to the eyes, ears, or nose. Internal administration may proceed intradermally, subcutaneously,

intramuscularly, intraperitoneally, intraarterially or intravenously, or by any other suitable route. It also may in some cases be advantageous to administer a composition of the invention by oral ingestion, by respiration, rectally, or vaginally.

The custom-made meganucleases and vectors can be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringers solution, and isotonic sodium chloride solution. Typically, for therapeutic applications, the custom-made meganucleases will be combined with a pharmaceutically acceptable excipient appropriate to a planned route of administration. A variety of pharmaceuticaly acceptable excipients are well known, from which those that are effective for delivering meganucleases to a site of infection may be selected. The HANDBOOK OF PHARMACEUTICAL EXCIPIENTS published by the American Pharmaceutical Association is one useful guide to appropriate excipients for use in the invention. A composition is said to be a "pharmaceutically acceptable excipient" if its administration can be tolerated by the recipient. Sterile phosphate-buffer saline is one example of pharmaceutically acceptable excipient that is appropriatae for intravenous administration.The mode of administration is preferably at the location of the targeted cells.

The dosage of custom-made meganuclease or vector of the present invention administered to an individual, including frequency of administration, will vary depending upon a variety of factors, including mode and route of administration: size, age, sex, health, body weight and diet of the recipient; nature and extent of symptoms of the disease or disorder being treated; kind of concurrent treatment, frequency of treatment, and the effect desired. For a brief review of pharmaceutical dosage forms and their use, see PHARMACEUTICAL DOSAGE FORMS AND THEIR USE (1985) (Hanshuber Publishers, Berne, Switzerland).

For purposes of therapy, the custom-made meganucleases and a pharmaceutically acceptable excipient are administered in a therapeutically effective amount. Such a combination is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of the recipient. In the present context, an agent is physiologically

significant if its presence results in a decrease in the severity of one or more symptoms of the targeted disease or in a genome correction of the lesion or abnormality.

In one embodiment of the invention, the custom-made meganucleases are substantially non-immunogenic, i.e., engender little or no adverse immunological response. A variety of methods for ameliorating or eliminating deleterious immunological reactions oft his sort can be used in accordance with the invention. In a preferred embodiment, the meganucleases are substantially free of N-formylmethionine. Another way to avoid unwanted immunological reactions is to conjugate meganucleases to polyethylene glycol ("PEG") or polypropylene glycol ("PPG") (preferably of 500 to 20,000 daltons average molecular weight (MW). Conjugation with PEG or PPG, as described by Davis et al., (US4,179,337) for example, can provide non-immunogenic, physiologically active, water soluble endonuclease conjugates with anti-viral activity. Similar methods also using a polyethylene--polypropylene glycol copolymer are described in Saifer et al. (US 5,006,333).

In another use, the custom-made meganucleases can be for *in vivo* excision of a polynucleotide fragment flanked by at least one, preferably.two, recognition and cleavage site for one or two distinct custom-mademeganucleases. Custom-made meganucteases according to the present invention can be used in methods involving the excision of targeted DNA or polynucleotide fragment from a vector within cells which have taken up the vector. Such methods involve the use of a vector comprising said poiynucleotide fragment flanked by at least one, preferably two, recognition and cleavage site for a custom-made meganuclease and either an expression vector comprising a polynucleotide encoding said custom-made meganuclease corresponding to the target site suitable for the expression in the used cel l, or said custom-made meganuclease. Said excised polynucleotide fragment can be used for transgenesis as described in detail in US patent application under number 10/242,664 filed on 13/09/02. Optionally, said excised targeting DNA comprises shared DNA homologies located in regions flanking upstream and downstream the site of the break in the targeting DNA construct and the DNA that might be introduced

being located between the two arms. For more detail see WO 00/46385. Said method of excision of targeting DNA from a vector within the cell can be used for repairing a specific sequence of interest in chromosomal DNA, for modifying a specific sequence or a gene in chromosomal DNA, for attenuating an endogeneous gene of interest, for introducing a mutation in a target site, or for treating or prophylaxis of a genetic disease in an individual.

- Antiviral application

There are currently very few effective anti-viral agents, although the many virally transmitted diseases account for much human suffering and mortality. Thus, there is a great need for safe and effective antiviral agents that can serve in therapies for these diseases, including life-threatening and fatal diseases such as hepatitis B and AIDS. 15 % of the human cancer have viral causes. This rate increases to 80 % for uterine and liver cancer. Today, viruses are the third carcinogenic factors in human.

No existing viral treatment is dedicated to kill the virus infecting the cells and cure the cells out of such infection.The main goal in treating viral infection is reducing viral load in infected cells and within a patient. Anti-viral drugs available today are generally toxic and have little specificity. Certain drugs are designed to inhibit a component of the virus's replicative machinery such as the enzymes thymidine kinase or reverse transcriptase. These agents do not destroy viral DNA. Other anti-viral agents act to promote the host's immune response so that infected cells are killed more efficiently. This results in non-specific destruction of both the virus and the host cell.

Today, there is a need for new therapeutic agents that specifically destroy viral DNA without destroying or altering the host cell. Most viral DNA synthesis occurs within the cell's nucleus; thus it is important to generate therapeutic agents that can distinguish between the viral and the host cell DNA.

Sechler (US 5,523,232) disclosed the use et restriction endonucleases against viruses. This method comprises a step of administering to a patient a composition with a restriction endonuclease able te cleave the targeted virus. The problem of this method is the very low specificity of the restriction endonuclease

(recognition sites of 4 to 6 nucleotides length) and the high probability to cleave the patient's genome.

Chandrasegaran (US 6,265,196) disclosed a prospective example concerning the use of a hybrid endonuclease (FoK I domain and zinc finger DNA binding domain) in treatment of viral diseases. However, such hybrid endonucleases generally lack the ability to specifically act as a single unique phosphodiester bond or base pair within the DNA target site (Smith et al., 1999, supra) and they do not show a high sequence specificity or any antiviral evidence.

One application of the custom-made meganucleases according to the present invention is as therapeutics in the treatment of viral diseases caused by viruses or retroviruses that present a DNA intermediate. Indeed, many viruses which infect eukaryotic cells possess, during at least one part of their life cycle, genomes that consist of double stranded DNA which can be cleaved readily by a meganuclease. Custom-made meganucleases according to the present invention can be designed so that they specifically target viral-specific DNA sequences.

At the opposite of restriction endonucleases and zinc fingers hybrid endonucleases, custom-made meganucleases according to the present invention present a good specificity for its viral DNA target and they generate a unique double strand break only in its viral target. This strategy involves identification of DNA sequences within the viral genome that are viral-specific, i.e., they are not present within the human genome. Once identified, meganucleases that specifically bind and cleave such sequences with high affinity and specificity can be designed using the method for preparing custom-made meganucleases as described in the present invention. Then the designed meganucleases are used for the treatment of viral infection.

It is therefore an object of the present invention to use at least one custom-made meganuclease according to the present invention to prevent, ameliorate or cure viral infection by exposing the virus and/or infected cells, plants, animals or patients to these meganucleases. Preferably the invention relates to the use of one meganuclease or two different maganucleases.

Another object of the present invention is to use at least one meganuclease according to the present invention for inactivating or deleting virus in biologically derived products and products intended for biological uses by treating the products with said meganucleases. in a particular embodiment, said biological products are blood or blood-derived products. Preferably the invention relates to the use of at least one meganuclease; optionally two different custom-made meganucleases.

Another object of the invention is a composition comprising at least one custom-made meganuclease according to the present invention for preventing, ameliorating or curing viral infection by exposing the virus or the infected cells, plants, animals or patients to said composition. Preferably, said composition comprises at least one custom-made meganuclease, optionally two meganucleases.

Another object of the invention, is compositions comprising at least one meganuclease according to the invention for inhibiting propagation of a virus, inactivating or deleting virus in biologically derived products or products intended for biological use, or for disinfecting an object. In a particular embodiment, said biological products are blood or blood-derived products. Preferably, said composition comprises one custom-made meganuclease or two different custom-made meganucleases

Any virus that contains a double stranded DNA stage in its life cycle can be targeted for deletion or inactivation by creating a meganuclease that recognizes DNA sequences specific of the viral genome. These viruses could be in replicative or latent form. They could stay either episomal or integrated in the host's genome.

The double stranded DNA genome viruses are well appropriate to be treated by using meganuclases as defined in the present invention. Among them are found the adenoviruses, the herpesviruses, the hepadnaviruses, the papovaviruses, and the poxviruses. Among the herpesviruses are found herpes simplex virus (HSV), varicella virus (VZV), Epstein-Barr virus (EBV), cytomegalo virus (CMV), herpes virus 6, 7 and 8. Among the hepadnaviruses are found the human hepatitis B virus (HBV). Among the papovariruses are found papillomavirus (HPV) (i.e. HPV16 or HPV18) and polyoma virus. Among the adenoviruses are found adenovirus 11 and 21

which are involved in acute hemorrhagic cystitis. Plants viruses are also contemplated by the present invention.

The retroviruses are also well appropriate to be treated by using meganucleases according to the present invention. Although they are RNA viruses, they are integrated in the host genome as double-stranded DNA form. Among the retroviruses are found the human immunodeficiency virus (HIV) and the human T lymphoma virus (HTLV) (i.e. HTLV1).

Several above-mentioned viruses are well-known to be involved in carcinogenesis: EBV in Burkitt's lymphoma, other lymphoproliferative disease and nasopharyngeal carcinoma; herpes virus 8 in Kaposi sarcoma; HBV in hepatocellular carcinoma; HPV in genital cancer; HTLV-1 in T-cell leukemia.

For episomal viruses, a double-strand break introduced in its genome leads to the linearisation of the genome and its degradation. Examples of episomal viruses are HSV-1, EBV, and HPV. See example 3.

For integrated viruses, a double strand break introduced in or near the integrated viral sequence leads to partial or complete deletion of the integrated viral sequence. Examples of integrated viruses are HPV, HTLV, HBV, and HIV. Several mechanisms could be involved in the deletion. A double-strand break in a chromosome induces a gene conversion with the homologous chromosome, therefore leading to viral sequence deletion. If directed repeat sequences are present near the double strand break, the break could also be repaired by SSA (single strand annealing) leading to partial or complete viral deletion. If two double-strand breaks are introduced, then the chromosome could also be repaired by end joining leading to partial or complete deletion of the virus, depending on the positions of the double-strand breaks. See Example 5 in US Patent n° 5,948,678, the disclosure of which is incorporated herein by reference.

To ensure that the targeted viral DNA sequences are not present in the host's genome, such DNA target sequences should be at least 15 nucleotides in length and preferably at least 18 nucleotides in length. As the homing endonuclease present a recognition sequence spanning to 12-40 bp, this condition is fulfilled with the custom-made meganucleases as defined in the present invention. More particularly, *I-Cre* I homing endonuclease has a 22 bp recognition sequence.

Any DNA sequence of viral genomes can be targeted for cleavage by meganucleases as defined in the present invention. Preferred target sites include those sequences that are conserved between strains of virus and/or which genes are essential for virus propagation or infectivity. These positions are preferable for at least two reasons. First, essential parts of viruses are less mutated than others. Secondly, it is preferably to target an essential region of the virus to maximize the inactivation of the virus.

A good target for the custom-made meganuclease could be the viral origin of replication (ori) and/or the viral gene encoding an ori binding protein. Examples of ori binding proteins include the HSV-1 UL9 gene product, the VZV gene 51 product, the human herpesvirus 6B CH6R gene product, the EBV EBNA-1 gene product and the HPV E1 and E2 gene products. Other interesting targets for HPV are the genes E6 and E7 as products of which are involved in the initiation and maintenance of the proliferative and malignant phenotype. A preferred target is the highly conserved 62 nucleotides sequence in the pre-core/core region of HPV (E6, E7). Examples of interesting targets for EBV are the genes EBNA and LMP. It could be interesting to target the gene Tax of HTLV-1 which appears to mediate the oncogenic effects of the virus. For HBV, an interesting target could be the X gene as the X protein interacts with elements of the DNA repair system and may increase the mutation rate of p53. For HIV, a preferred target is within TAT, REV, or TAR genes. The viral targets are not limited to the above-mentioned examples. Optionally, the target DNA could be located in the viral repeated sequences such as ITR (Inverted Terminal Repeat) and LTR (Long Terminal Repeat).

Preferably, at least two different targeted sites are used. Indeed, as the main protection of the viruses is their ability to mutate. Therefore, two targeted sites avoid the virus to escape the treatment by using the custom-made meganucleases, according to the present invention. Moreover, the successive use of different custom-made meganucleases may avoid the adverse immunologic response. Said different custom-made meganuclease can present different initial meganucleases, therefore different immunogenicities.

The treatment by custom-made meganucleases according to the present invention can be applied either on cells, preferably cells taken from the patient, or on the whole body of the patient, with or without organ targeting.

In the case of cell therapy, cells are preferably taken from a patient. These cells are treated by custom-made meganucleases according te the present invention in order to inactivate oo delete the virus. After the treatment, cells are reintroduced into the patient. These cells will proliferate and repopulate infected tissues. Preferably, said cells are stem cells, totipotent cells or pluripotent cells. For example, stem cells could be hematopoietic, neuronal, mesenchymal, embryonic, muscle-derived. Another examples of cells are those which are able te regenerate such as the hepatocytes. The treatment by custom-made meganucleases can be done either by the introduction of meganucleases into cells or by the transfection with an expression vector encoding said meganucleases. The transfection can be transient or stable. A transient expression of the custom-made meganuclease allows the cell to be cleaned up from the virus. A stable transfection allows the cell to be cleaned up from the virus and avoids a further infection of the treated cells by the targeted virus.

In the case of a whole therapy, custom-made meganucleases according ta the present invention or expression vector encoding said meganucleases are introduced into the individual by any convenient mean. When the custom-made meganucleases are introduced or expressed into the infected cells, the virus is inactivated and/or deleted. The custom-made meganuclease treatment has no functional impact on healthy cells.

Similarly, such cell or whole antiviral therapy based on the use of at least one custom-made meganuclease could be used to treat cells or organs of an animal dedicated to xenotransplantation. The effectiveness of a meganuclease to inhibit viral propagation and infection is preferably assessed by *in vitro* and *in vivo* assays of infection. Such assays can be carried out first in cell culture to establish the potential of different meganucleases to cleave a viral DNA in a way that deleteriously affects viral propagation. Preliminary studies of this type are followed by studies in appropriate animal models. Finally, clinical studies will be carried out.

Different viruses require different assay systems, since hosts and culture conditions suitable to different viruses vary greatly. However, such appropriate conditions have been described for culturing many viruses and these conditions can be used to test the effect of exposing virus and/or host to meganucleases to determine the ability of the endonuclease to inhibit viral infection. For one discussion of culture conditions for specific viruses see Chapter 17 in Fields and Knipe, Eds., FIELDS VIROLOGY, 2nd Ed., Raven Press, N.Y. (1990).

A host and/or virus can be exposed at various times during a course of infection, under varying conditions, in several amounts, and in a variety of vehicles, to mention just a few relevant parameters that can be varied, to assess the potential of meganuclease to achieve a potentially therapeutic effect.

In addition, in order to tests *ex vivo* in cultured cells, potential therapeutical meganuclease can be tested in animal models to assess prophylactic, ameliorative, therapeutic and/or curative potential, either alone or in conjunction with other therapeutic agents. In some cases, it will not be possible to culture a virus and it will be necessary to perform all biological assays in animal models. It will be readily appreciated that different animal models will be appropriate to different viruses. Any animal model, however, can be used to assess the therapeutic potential of a meganuclease.

A potentially effective dose of the assayed meganucleases may be administered to a suitable population of animals, and the effect of the meganucleases on the course of a viral infection may be assessed by comparison with an appropriate control. Such methods for assessing pharmacological effect are well known in the art and can readily be adapted to determining the therapeutic profile of the meganucleases.

In one embodiment of the uses according to the present invention, the meganuclease is substantially non-immunogenic, i.e., engender little or no adverse immunological response. A variety of methods for ameliorating or eliminating deleterious immunological reactions of this sort can be used in accordance with the invention. In a preferred embodiment, the meganuclease is substantially free of N-formyl methionine. Another way to avoid unwanted immunological reactions is to conjugate meganucleases to polyethylene glycol ("PEG") or polypropylene glycol ("PPG")

(preferably of 500 to 20,000 daltons average molecular weight (MW)). Conjugation with PEG or PPG, as described by Davis et al., (US 4,179,337) for example, can provide non-immunogenic, physiologically active, water soluble endonuclease conjugates with anti-viral activity. Similar methods also using a polyethylene--polypropylene glycol copolymer are described in Saifer et al. (US 5,006,333).

Custom-made meganuclease according to the present invention can be introduced into cells using liposomes or by fusion to the membrane translocating peptides (Bonetta 2002, The Scientist, 16, 38; Ford et al, Gene Ther, 2001, 8, 1-4; Wadia & Dowdy Curr Opin Biotechnol, 13, 52-56). Otherwise, meganucleases can be stably or transiently expressed into cells using expression vectors. Techniques of expression in eukaryotic cells are well known to those in the art. (See Current Protocols in Human Genetics: Chapter 12 "Vectors For Gene Therapy" & Chapter 13 "Delivery Systems for Gene Therapy"). Optionally, it may be preferable to incorporate a nuclear localization signal into the recombinant protein to be sure that it is expressed within the nucleus.

Typically, for therapeutic applications, the custom-made meganucleases will be combined with a pharmaceuticaliy acceptable excipient appropriate to a planned route of administration. A variety of pharmaceuticaIly acceptable excipients are well known, from which those that are effective for delivering meganucleases to a site af infection may be selected. The HANDBOOK 0F PHARMACEUTICAL EXCIPIENTS published by the American Pharmaceutical Association is one useful guide to appropriate excipients for use in the invention. A composition is said ta be a "pharmaceutically acceptable excipient" if its administration can be tolerated by the recipient. Sterile phosphate-buffered saline is one example of a pharmaceutically acceptable excipient that is appropriate for intravenous administration.

For purposes of therapy, the custom-made meganucleases and a pharmaceutically acceptable excipient are administered in a therapeuticaily effective amount. Said composition can comprise either one kind of custom-made meganuclease or several custom-made meganucleases with different specificity. Such a combination is said to be administered in a "therapeutically effective amount" if the

amount administered is physioiogically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of the recipient. In the present context, an agent is physiologically significant if its presence results in a decrease in the severity of one or more symptoms of a viral illness.

Administration may be topical or internal or other suitable avenue for introducing a therapeutic agent to a patient. Topical administration may be by application to the skin, or to the eyes, ears or nose. Internal administration may proceed intradermally, subcutaneously, intraperitoneally, intraarterially or intravenously, or by any other suitable route. It also may in some cases be advantageous to administer a composition of the invention by oral ingestion, by respiration, rectally, or vaginally. For a brief review of pharmaceutical dosage forms and their use, see PHARMACEUTICAL DOSAGE FORMS AND THEIR USE (1985) (Hans Huber Publishers, Berne, Switzerland).

For topical and internal therapeutic applications, custom-made meganucleases according to the present invention can be formulated using any suitable pharmacological technique. For instance, the meganucleases can be formulated for prolonged release. As described hereinabove, persistence of anti-viral activity of the meganucteases may be increased or modulated by incorporating the meganucleases into liposomes.

Additionally, the method described in the present invention could be used to modify the physico-chemical properties of meganucleases. For example, said method could be used to change the sensitivity of a meganuclease to temperature, pH or salt concentration (e.g. in order to decrease or increase the level at which activity is highest or to enhance the activity at a chosen level), as well as its solubility or stability, or its reaction turnover. Eventually, the method described in the present invention could also be used to relax or strengthen the specificity of given meganucleases for preferred DNA sequence targets.

The present invention will be further illustrated by the additional description and drawings which follows, which refers to examples illustrating the method to produce a custom-made according to the invention and the use thereof for antiviral therapy and gene therapy. It should be understood however that these

examples are given only by way of illustration of the invention and do not constitute in anyway a limitation thereof.

- Figure 1 discloses the amino acid sequence of a single chain I-Cre I meganuclease and one polynucleotide encoding said single chain meganuclease. In the protein sequence, the two first N-terminal residues are methionine and alanine (MA), and the three C-terminal residues alanine, alanine and aspartic acid (AAD). These sequences allow having DNA coding sequences comprising the NcoI (CCATGG) and EagI (CGGCCG) restriction sites, which are used for cloning into various vectors.

- Figure 2 discloses polynucleotide sequences. Figure 2A discloses a polynucleotide called "Natural" encoding the I-Cre I homing endonuclease. Figure 2B discloses a polynucleotide sequence called "Non homologous" encoding the I-Cre I homing endonuclease. Figure 2C discloses a polynucleotide sequence called "Template" encoding the I-Cre I homing endonuclease comprising the mutation D75NJ. Each I-Cre I homing endonuclease has two additional amino acids (MA) at the N terminal end and three additional amino acids (AAD) at the C-terminal ends. Figure 2D discloses the polynucleotide sequences of the primers, called UlibIfor, UlibIrev, UlibIIfor, and UlibIIrev, used for the generation of the libraries UlibI and UlibII.

- Figure 3 is a schematic representation of the polynucleotide sequence called "Template" encoding the I-Cre I homing endonuclease comprising the mutation D75N. The dark arrows indicate the position of the primers UlibIfor, UlibIrev, UlibIIfor, and UlibIIrev used to generate the two libraries UlibI and UliblII. D-Helix refers to the LAGLIDADG helix. N75 refers to the mutation D75N.

- Figure 4 is a schematic representation of the strategy for the library construction. Step 1: pET24C-T is a plasmid comprising a polynucleotide «Template». Two PCR amplifications, PCR ulib1 and ulib2, are done with either UlibIfor and UlibIIrev, or UliblIfor and UliblIIrev. The PCR ulib1 products are cloned in a phagemid pCes4 NHT. The PCR ulib2 products are cloned in a plasmid pET24C-T. Step 2: Subcloning a fragment of Ulib2 vector (pET45C-Ulib2) into the Ulib1 phagemid (pCes4-Ulib1).

- Figure 5 : COS cells monolayers were transfected with vector expressing I Sce-I (B) or with control plasmid (A). Fourty eight (48) hours after

transfection cells were infected with rHSV-1 (30 PFU). Two days later monolayer was fixed and stained (X-Gal). Infected cells appeared in blue.

- Figure 6 : figure 6A: cells monolayer was infected with 30 PFU. HSV-1 growth was quantified by ß-galactosidase activity in cell lysate. Figure 6B, cell monolayer was infected with 300 PFU. Cell survival was measured by protein determination in cell lysate. I-Sce I refers to vector expressing I-Sce I ; I-Sce I(-) refers to a vector in which ORF of I Sce-I was inserted in reverse orientation ; negative control refers to control plasmid.

- Figure 7 : figure 7A is a schematic representation of recombinant HSV-1 genomic DNA. Cassette containing CMV promoter driving Lac gene was inserted in the major LAT transcript. I-Sce I restriction site was cloned between promoter and reporter gene. a and b represent primers used for the semi-quantitative PCR. COS-7 monolayers were transfected with vector expressing I-Sce I or with control plasmids. Fourty eight hours after transfection cells were infected with rHSV-1 (30 PFU). DNA was extracted 1, 2 or 3 days after infection. PCR was carried out as described in « experimental procedures ». Std refers to Internal standard; Lac refers to an amplicon of the rHSV-1 Lac gene. I-Sce I refers to vector expressing I-Sce I ; I-Sce I(-) refers to a vector in which ORF of I-Sce I was inserted in reverse orientation ; negative control refers to control plasmid. Figure 7B, PCR quantification of the viral thymidine kinase (TK) gene. PCR was carried out at 2 DNA concentrations. ISce-I refers to vector expressing I-Sce I; I-Sce I(-) refers to a vector in which ORF of I-Sce I was inserted in reverse orientation ; negative control refers to control plasmid.

- Figure 8 illustrates the titration of the virus released in the medium after infection of the transfected cells. Every day, medium was collected and fresh medium was added. Viruses were measured by standard plaque assay. I-Sce I refers to vector expressing I-SceI ; I-Sce I(-) refers to a vector in which ORF of I-Sce I was inserted in reverse orientation ; negative control refers to control plasmid.

- Figure 9 represents the I-*Cre*I DNA target and five related targets. Conserved positions are in grey boxes.

- Figure 10 illustrates four binding patterns obtained after screening of the Lib2 library with six targets. Positives were identified in a first screen and confirmed in a second one during which they were assayed eight times (corresponding

to the eight solid bars) on each of the targets (C1234, C1221, C4334, H1234, H1221 and H4334). Histograms are shown for one clone from each class. Targets are described in Figure 9.

- Figure 11 illustrates the schematic representation of the target vectors. The CpG depleted LacZ gene (*LagoZ*) is driven by the human elongation factor 1 alpha promoter. The *LagoZ* gene is inactivated by the insertion of *I-SceI* cleavage site. Flanking repeats are represented by open arrows. The length of the homologous sequences are indicated in bold.

- Figure 12 illustrates the effect of the length of homology on single strand annealing (SSA) efficiency. Cells monolayers were transfected with equimolar amounts of target plasmid bearing different lengths of homologous repeat sequences and vector expressing ISce-I or with control plasmid. Seventy-two hours after transfection cells were collected and ß-galactosidase activity was quantified in cell lysates. (+)*I-SceI*, cotransfection with vector expressing *I-SceI*; (-)*I-SceI*, cotransfection with expression vector where the ORF of *I-SceI* was inserted in the reverse orientation .

- Figure 13: Cell monolayers were cotransfected with a vector expressing (+)*I-SceI* or with a control plasmid (-)*I-SceI*. Seventy-two hours after transfection cells were fixed and stained (X-Gal). Figure 13A : cells where gene repair took place appeared in dark. Figure 13B : frequency of *I-SceI* induced recombination on 70 and 220 bp duplication target vectors. The frequency is calculated by the ratio of blue cells/transfected cells.

- Figure 14A : X-Gal staining of liver from mice injected with a mixture of the target LagoZ gene (30 µg) and an I-*Sce*I expression vector (10 µg). Figure 14B : X-Gal staining of liver from mice injected with a mixture of the target LagoZ gene (30 µg) and an expression vector where the ORF of I-*Sce*I was inserted in the reverse orientation (10 µg).

- Figure 15 : X-Gal staining of the liver of hemizygote transgenic mice of two independent strains infected with the « Ad.I-*Sce*I » adenovirus by IV. **A.** Five days post-infection, ß-galactosidase activity is detected in multiple cells of the entire liver of 10^{10} infectious units infected « 58A » hemizygote. In contrast, no β-galactosidase activity could be detected by X-Gal staining of the livers of « Ad.control »-infected hemizygote or un-infected « 58A » littermates (data not

shown). **B.** and **C.** Fourteen days post-infection, β-galactosidase activity is detected in multiple cells of the entire liver of 10^9 infectious units infected mouse (B) and 10^{10} infectious units infected mouse (C). Stronger signal is detected in C compared to B, probably because of the bigger number of cells that were infected with the «Ad.I-*Sce*I». In contrast, no ß-galactosidase activity could be detected by X-Gal staining of the livers of un-infected « 361 » littermates (data not shown).

- Figure 16 : Fluorescent ß-galactosidase assay on liver extract. Two independent strains of transgenic mice (58 A and 361) were injected with 10^9 or 10^{10} PFU of adenovirus expressing I-*Sce*I (Ad.I-*Sce*I) or control virus (Ad.control). Mice were sacrified 5 or 14 days post injection, liver was dissected and proteins were extracted. 30 µl of liver protein extract were incubated at 37°C in presence of Fluorescein digalactoside (FDG). Bars represent the standard deviation of the assay (two measure experiments with samples of the same extracts). NI, non injected mice; Ad.I-*Sce*I, mice injected with adenovirus expressing I-SceI ; Ad.control, mice injected with control adenovirus.

Example 1: Single chain meganuclease derived from dimeric homing endonucleases.

Some LAGLIDADG homing endonucleases are active as homodimer. Each monomer mainly dimerizes through their dodecapeptide motifs. A single-chain meganuclease can be engineered by covalently binding two monomers modified such as to introduce a covalent link between the two sub-units of this enzyme. Preferably, the covalent link is introduced by creating a peptide bond between the two monomers. However, other convenient covalent links are also contemplated. The single-chain meganuclease preferably comprises two subunits from the same homing endonuclease such as single-chain I-Cre I and single-chain I-Ceu I. A single-chain meganuclease has multiple advantages. For example, a single-chain meganuclease is easier to manipulate. The single-chain meganuclease is thermo-dynamically favored, for example for the recognition of the target sequence, compared to a dimer formation. The single-chain meganuclease allows the control of the oligomerisation.

A single chain version of I-CreI (scI-CreI) was modeled and engineered. scI-CreI cleaves its cognate DNA substrate *in vitro* and induces homologous recombination both in yeast and mammalian cells.

- Design of the single chain I-CreI meganuclease

I-CreI from *Chlamydomonas reinhardtii* is a small LAGLIDADG homing endonuclease that dimerizes into a structure similar to that of larger monomer LAGLIDADG homing endonuclease. To engineer a single chain version of I-CreI (scI-CreI), two I-CreI copies were fused. This required placing a linker region between the two domains, and a significant part of the I-CreI protein had to be removed at the end of the domain preceding the linker.

The three-dimensional structure of I-DmoI is comparable to that of I-CreI, with the exception that I-DmoI comprises a linker region that leads from one apparent domain to the other. The boundary of that linker finely matches related main chain atoms of the I-CreI dimer. In the first domain, residues 93 to 95 from the third α-helices of I-CreI and I-DmoI (prior to the linker) are structurally equivalent. At the beginning of the second LAGLIDADG α-helix (second domain), I-DmoI residues 104 to 106 correspond to I-CreI residues 7 to 9. In addition, Leu95 and Glu105 from I-DmoI have conserved identities in I-CreI, and I-DmoI residue Arg104 aligns with another basic residue in I-CreI (Lys7). Thus, the single chain I-CreI (scI-CreI), was designed by inserting the I-DmoI linker region from residue 94 to 104 (sequence MLERIRLFNMR) between a first I-CreI domain (terminated at Pro93) and a second I-CreI domain (starting at Glu8).

Detailed structural analysis of how the new linker connects the scI-CreI protein domains (in a modeled structure) revealed no potential incompatibility. For example, the side chains of nonpolar amino acids taken from I-DmoI, Met94, Ile98 and Phe109 point inside fitting cavities of I-CreI. A single mutation was made (P93A), however, to promote regularity of the backbone in the α-helix prior to the linker region. (See Figure 1 for amino acids and polynucleotide sequences).

- Materials and methods

Protein expression and purification

His-tagged proteins were over-expressed in *E. coli* BL21 (DE3) cells using pET-24d (+) vectors (Novagen). Induction with IPTG (1mM), was performed at

25°C. Cells were sonicated in a solution of 25mM HEPES (pH 8) containing protease inhibitors (Complete EDTA-free tablets, Roche) and 5% (v/v) glycerol. Cell lysates were centrifuged twice (15 000 g for 30 min). His-tagged proteins were then affinity-purified, using 5ml Hi-Trap chelating columns (Amersham) loaded with cobalt. Several fractions were collected during elution with a linear gradient of immidazole (up to 0.25M immidazole, followed by plateau at 0.5M immidazole and 0.5M NaCl). Protein-rich fractions (determined by SDS-PAGE) were concentrated with a 10kDa cut-off centriprep Amicon system. The resulting sample was eventually purified by exclusion chromatography on a Superdex75 PG Hi-Load 26-60 column (Amersham). Fractions collected were submitted to SDS-PAGE. Selected protein fractions concentrated and dialyzed against a solution of 25mM HEPES (pH 7.5) and 20% (v/v) glycerol.

In vitro cleavage assays

pGEM plasmids with single meganuclease DNA target cut sites were first linearized with XmnI. Cleavage assays were performed at 37°C or 65°C in 12.5mM HEPES (pH 8), 2.5% (v/v) glycerol and 10mM MgCl2. Reactions were stopped by addition of 0.1 volume of 0.1 M Tris–HCl (pH 7.5), 0.25 M EDTA, 5% (w/v) SDS, and 0.5 mg/ml proteinase K and incubation at 37°C for 20 minutes. Reaction products were examined following separation by electrophoresis in 1% agarose gels.

Yeast colorimetric assay.

The yeast transformation method has been adapted from previous protocols. For staining, a classic qualitative X-Gal Agarose Overlay Assay was used. Each plate was covered with 2.5 ml of 1% agarose in 0.1 M Sodium Phosphate buffer, pH 7.0, 0.2% SDS, 12% Dimethyl Formamide (DMF), 14 mM β-mercaptoethanol, 0.4% X-Gal, at 60°. Plates were incubated at 37°C.

Mammalian cells assays

COS cells were transfected with Superfect transfection reagent accordingly to the supplier (Qiagen) protocol. 72 hours after transfection, cells were rinsed twice with PBS1X and incubated in lysis buffer (Tris-HCl 10mM pH7.5, NaCl 150mM, Triton X100 0,1 %, BSA 0.1 mg/ml, protease inhibitors). Lysate was centrifuged and the supernatant used for protein concentration determination and β-

galactosidase liquid assay. Typically, 30 µl of extract were combined with 3 µl Mg 100X buffer ($MgCl_2$ 100mM, β-mercaptoethanol 35%), 33µl ONPG 8 mg/ml and 234µl sodium phosphate 0.1M pH7.5. After incubation at 37°C, the reaction was stopped with 500µl of 1M Na_2CO_3 and OD was measured at 415nm. The relative β-galactosidase activity is determined as a function of this OD, normalized by the reaction time, and the total protein quantity.

- Results : Single chain I-CreI cleaves its DNA substrate *in vitro* and in living cells

A synthetic gene corresponding to the new enzyme was engineered and the scI-CreI protein over-expressed in *E. coli*. The ability of purified scI-CreI to cleave DNA substrates *in vitro* was tested, using linearized plasmids bearing a copy of the I-CreI homing site. Similarly to parent I-CreI, the novel enzyme cleaves an I-CreI target site at 37°C.

In order to test the functionality of scI-CreI *in vivo*, an assay to monitor meganuclease-induced homologous recombination in yeast and mammalian cells was designed. In yeast, Xenopus oocytes and mammalian cells, DNA cleavage between two direct repeats is known to induce a very high level of homologous recombination between the repeats. The recombination pathway, often referred to as Single-Strand Annealing (SSA), removes one repeat unit and all intervening sequences. Thus, a SSA reporter vector, with two truncated, non-functional copies of the bacterial LacZ gene and an I-CreI cut site within the intervening sequence was constructed in a yeast replicative plasmid. Cleavage of the cut site should result in a unique, functional LacZ copy that can be easily detected by X-gal staining.

The reporter vector was used to transform yeast cells. A small fraction of cells appeared to express functional LacZ, probably due to recombination events during transformation. Co-transformation with plasmids expressing either I-CreI or scI-CreI, in contrast, resulted in blue staining for all plated cells. Even in non-induced conditions (glucose), the residual level of protein was enough to induce SSA, suggesting that scI-CreI, as much as I-CreI, is highly efficient in yeast cells. Furthermore, SSA induction was truly dependent on cleavage of the target cut site by I-CreI proteins, as vectors devoid of that site display no increase in β-galactosidase activity compared to background levels.

The SSA assay was modified for tests in mammalian cells. The promoter and termination sequences of the reporter and meganuclease expression plasmid were changed, and plasmid recombination was evaluated in a transient transfection assay. Similar levels of induced recombination (2 to 3-fold increase) were observed with either scI-CreI or I-CreI. As in the yeast experiment, recombination depends on an I-CreI cut site between the repeats, for no increase of the β-galactosidase was observed in the absence of this site.

Another recombination assay, based on recombination between inverted repeats, was also used to monitor meganuclease-induced recombination in COS cells. As direct repeats can recombine by SSA, homologous recombination between indirect repeats requires a gene conversion event. Similar stimulation of gene conversion (3 to 4- fold) was observed with either scI-CreI or I-CreI. As expected for a true homologous recombination event, no enhancement was observed in the absence of an homologous donor template.

Example 2: Custom-made meganuclease derived from I-Cre I homing endonuclease for HIV-2 target.

- Construction of a phage-displayed library of I-Cre I variants

In order to engineer new meganuclease with altered specificities, a combinatorial library was constructed by mutagenesis of the I-Cre I homing endonuclease replacing DNA binding residues. Selection and screening applications then enabled to find those variants that were able to bind a particular, chosen DNA target. For phage display, as I-Cre I is a homodimer, a phagemid vector was required that encoded two separate I-Cre I proteins. Only one of the two I-Cre I copies, which was fused to the phage coat protein p3, was mutated. The resulting protein library, in phage display format, comprised thus I-Cre I wild-type/mutant heterodimers. Eight residues (Q26, K28, N30, Y33, Q38, Q44, R68 and R70) capable together of specific interactions with most of the bases in a single hal-site within the DNA target were selected. Our combinatorial library was obtained by replacing the eight corresponding codons with a unique degenerated VVK codon. Eventually, mutants in the protein library corresponded to independant combinations of any of the 12 amino acids encoded by the VVK codon (ADEGHKNPQRST) at eight residue positions. In consequence, the maximal (theoretical) diversity of the protein library was 12^8 or $4.29x10^8$.

- Construction of the library

First, residue D75, which is shielded from solvent by R68 and R70, was mutated to N (Asn) in order to remove the likely energetic strain caused by replacements of those two basic residues in the library. Homodimers of mutant D75N (purified from *E. coli* cells wherein it was over-expressed using a pET expression vector) were shown to cleave the I-CreI homing site. A phagemid vector was then engineered that encodes wild-type I-CreI (Figure 2A and 2B: « Natural » or « Non homologous ») and the D75N mutant (Figure 2C: «Template») fused to the phage coat protein p3 and phage-displayed wild-type/D75N heterodimers were shown to bind that target DNA.

Second, two intermediate libraries of moderate size have been built: Lib1 (residues 26, 28, 30, 33 and 38 mutated; theoretical diversity 12^5 or $2.48x10^5$) and Lib2 (residues 44, 68 and 70 mutated; theoretical diversity 12^3 or $1,7x10^3$). DNA fragments carrying combinations of the desired mutations were obtained by PCR (several reactions in 50 µl), using degenerated primers (Figure 2D: Uliblfor, Uliblrev, Ulibllfor, Ulibllrev) and as DNA template, the D75N gene. Lib1 and Lib2 were constructed by ligation of the corresponding PCR products, digested with specific restriction enzymes, into the D75N mutant gene, within the phagemid vector and within the pET expression vector, respectively. Digestions of vectors and inserts DNA were conducted in two steps (single enzyme digestions) between which the DNA sample was extracted (phenol:chloroform:isoamylalcohol) and EtOH-precipitated. 10 µg of digested vector DNA were used for ligations, with a 5:1 excess of insert DNA. *E. coli* TG1 cells were transformed with the resulting vectors by electroporation. To produce a number of cell clones above the theoretical diversity of either library, up to 35 electroporations of the Lib1 ligation samples and 4 electroporations of the Lib2 ligation samples were necessary. $4x10^6$ (16 times the maximal diversity) and $6x10^4$ (35 times the diversity) clones were thus obtained for Lib1 and Lib2, respectively (these numbers were corrected by the number of clones obtained using ligations done without inserts).

Finally, Lib1 and Lib2 bacterial clones were scraped from plates and the corresponding plasmid vectors were extracted and purified. The complete library

was then obtained by sub-cloning a fragment of the Lib2 vector into the Lib1 phagemid vector (see Figure 4 for a schematic diagram of the library construction). Several rounds of DNA 2-step digestions, dephosphorylation, purification, quantification, ligation and electroporation were performed. After 4 rounds of 150 electroporation shots (which corresponds to 12 ligations of 1.4 µg vector with 0.4 µg insert), $5.5x10^7$ bacterial clones were obtained (after correction for background). Bacteria were scraped and stored as a glycerol stock. In addition, an aliquot of this glycerol stock was used to inoculate a 200 ml culture and the library vector was extracted and purified from this culture for storage or potential subcloning.

- Material and methods

Protein expression and purification

His-tagged proteins were over-expressed in *E. coli* BL21 (DE3) cells using pET 24d (+) vectors (Novagen). Induction with IPTG (1 mM), was performed at 15°C over 5 night. Cells were cracked for 1h at 4°C in a B-Per solution (Bacterial Protein Extraction Reagent, Pierce, 5ml for 200ml culture cell), containing protease inhibitors (Complete EDTA-free tablets, Roche) and DNase I (80 units) / nuclease (respectively 80 and 60 units, Roche). Alternatively, cells were sonicated in a solution of 25 mM HEPES (pH 8) containing protease inhibitors (Complete EDTA-free tablets, Roche) and 5% (v/v) glycerol.

Cell lysates were centrifuged twice (15 000 g for 30 min). His-tagged proteins were then affinity-purified, using 1 ml Hi-Trap chelating columns (Amersham) loaded with cobalt. Several fractions were collected during elution with a linear gradient of immidazole (up to 0.25 M immidazole, followed by plateau at 0.5 M immidazole and 0.5 M NaCI). Protein-rich fractions (determined by SDS-PAGE) were concentrated with a 10 kDa cut-off centriprep Amicon system. The resulting sample was eventually purified by exclusion chromatography on a Superdex75 PG Hi-Load 26-60 column (Amersham).

Fractions collected were submitted to SDS-PAGE. Selected protein fractions concentrated and dialyzed against a solution of 25 mM HEPES (pH 7.5) and 20% (v/v) glycerol.

In vitro cleavage assay

pGEM plasmids with single meganuclease DNA target cut sites were first linearized with XmnI. Cleavage assays were performed at 37°C in 12.5 mM HEPES (pH 8), 2.5% (v/v) glycerol and 10 mM $MgCl_2$. Reactions were stopped by addition of 0.1 volume of 0.1 M Tris—HCl (pH 7.5), 0.25 M EDTA, 5% (w/v) SDS, and 0.5 mg/ml proteinase K and incubation at 37°C for 20 minutes. Reaction products were examined following separation by electrophoresis in 1% agarose gels.

Phagemid construction

Phage Display of I-Cre I/D75N heterodimer was obtained by using a phagemid harboring two different ORFs as a bicistron, under the control of promoter pLac. The first one yields a soluble protein fused to a N-terminal signal sequence directing the product into the periplasmic space of *E. coli* Gene 1-Cre I WT was cloned into this ORF using restriction enzymes ApaLI and AscI. The D75N domain was cloned into the second ORF using Nco I and Eag I restriction enzyme, leading to a fusion with the phage coat protein p3 via a hexahis tag, a C-Myc tag and an amber stop codon. This final phagemid was called pCes1CreT. In a suppressive strain like TG1 or XL1blue, and after infection by a helper phage (e.g. M13K07), D75N-p3 fusions are incorporated in the phage coat and the soluble I-CreI mononers produced in the same compartment will either dimerize or interact with the displayed D75N domain, thereby producing particles displaying I-CreI WT/D75N heterodimer.

Phage production

A 5 mL culture of 2xTY containing 100 µg/ml of ampicillin and 2% glucose was inoculated with a 1/100 dilution of an overnight culture of bacteria containing phagemid pCes1CreT and agitated at 37°C. At an OD_{600} of 0.5, phage helper M13K07 (Pharmacia) was added at a ratio phage:bacteria of 20:1. After 30 min at 37°C without agitation, the culture was centrifuged for 10 min at 4000 rpm and the pellet was resuspended in 25 ml of 2xTY containing 100 µg/mL Ampicillin and 25 µg/mL Kanamycin, and agitated overnight at 30°C. Cultures were centrifuged and supernatant were used as such in phage ELISA.

PhageELISA

Microtiter plates were coated for 1h at 37°C with 100 µl/well of biotinylated BSA at 2 µg/mL in PBS. After several washes in PBS containing 0.1% Tween20 (PBST), wells were incubated with 100 µl/well of streptavidin at 10 µg/mL in PBS and incubated for 1h at RT. Plates were further washed and incubated with biotinylated PCR fragments harboring the target site, at 250 pM in PBS. After 1h incubation at RT and washing, plates were saturated with 200 µl/well of PBS containing 3% powder milk and 25 mM $CaCl_2$ (PMC). PMC was discarded and plates were filled with 80 µl of PMC and 20 µl/well of culture supernatant containing the phage particles. After 1h of incubation at RT, plates were extensively washed with PBST and incubated with 100 µl/well of anti 25 M13-HRP conjugated antibody (Pharmacia) diluted 1/5000 in PMC. Plates were incubated for 1h at RT, washed and incubated with TMB solution (Sigma). The reaction was blocked with 50 µl/well of 1M H_2SO_4. Plates were read at 450 nm. A signal higher than 3X the background (irrelevant target) can be considered as positive.

PCR-based mutagenesis

Plasmid pET24-T45 containing the gene I-CreI D75N was diluted at 1 ng/µl to be used as template for PCR. Degenerated oligonucleotides encoding the desired randomizations were used to amplify PCR fragments Lib1 and Lib2 in 4x50 µl PCR reactions per inserts. PCR products were pooled, EtOH precipitated and resuspended in 50 µl 10 mM Tris.

DNA digestions

All enzymes and the corresponding buffers were from NEBiolabs. Digestions of up to 10 µg DNA were realised using up to 100 U of a first restriction enzyme, at 37°C, in 150 or 500 µl final reaction volume. After 2h to 6h, digested DNA was phenol extracted and EtOH precipitated. Digestion substrates and products were separated using agarose gel electrophoresis, the desired product being extracted from the gel and purified (Nucleospin Extract, Macherey-Nagel). For PCR inserts, digestions were directly purified on Nucleospin columns. The second digestion was then performed in identical conditions. At the end of this second digestion reaction, 0.1 volume of 10X CAP buffer and 0.5 µl of CAP were added to the digested vectors,

and the samples were further incubated for 30 min at 37°C (The alkaline phosphatase was inactivated by incubating the sample 10 min at 70°C, after addition of EDTA). Eventually, the digested and de-phosphorylated DNA was phenol extracted, EtOH precipitated and resuspended in 30 µl of 10 mM Tris pH8. Final DNA concentrations were estimated by comparison of band intensities in agarose gels after electrophoresis.

Ligations

Large-scale ligations were done at 16°C (for 16 h) using 1400 ng of digested vector and a 5:1 molar excess of digested in 200 µl reaction volumes and with 4000 U of T4 DNA ligase (NEBiolabs). After ligation, reaction samples were incubated for 20 min at 65°C to inactivate the ligase. The vector DNA was eventually EtOH precipitated and resuspended at 25 ng/µl in 10 mM Tris pH8.

Electroporations

40 µl of homemade electrocompetent cells TG1 were mixed with 25 ng of ligated DNA (1 µl) in a 2 mm cuvette. After 1 min on ice, cells were pulsed (2.5 Kv, 25 µF, 200 Ohm) and immediately resuspended in 1 ml of 2xTY + 2% glucose. Cells were placed at 37°C for 1 h with agitation, and then plated on large 2xTY plates containing ampicillin (phagemid vector) or kanamycin (pET vector) and 2% glucose and incubated overnight at 30°C. Aliquots were also diluted in 2xTY and plates on small 2xTY Ampicillin glucose plates to obtain isolated colonies allowing the calculation of library diversities and characterization of several clones by restriction analysis.

- Selection and screening of meganuclease binding to a HIV2-derived DNA target from a library of I-Cre I variant using Phage Display

The goal of this project was to obtain a meganuclease capable of cutting a sequence found in the genome of HIV2 (GGAAGAAGCCTTAAGACATTTTGA). The homing endonuclease I-Cre I was used as a scaffold to build a library of 10^8 variants by randomizing 8 residues located at the DNA-binding interface of one I-Cre I monomer (see previous section). This library was enriched for binders by several rounds of selection /amplification using

biotinylated DNA fragments harboring the HIV2 derived target (H1V6335). The selected targets were subsequently screened for binding using a phage ELISA.

-Materials and Methods

Phagemid format

A phagemid based on pCesl (pCLS346) was chosen. This plasmid harbored two different ORFs as a bicistron, under the control of promoter pLac. The first one yielded a soluble protein fused to a N-terminal signal sequence directing the product into the periplasmic space of *E. coli*. In our case, this first product was a wild-type monomer of I-CreI. The second ORF encoded an I-CreI monomer that was fused to the phage coat protein p3 via a hexahis tag, a C-Myc tag and an amber stop codon. In a suppressive strain like TG1 or XL1blue, and after infection by a helper phage (e.g. M13K07), bacteria harboring this phagemid produces phage particles and around 1-10% of them displays the recombinant protein on their surface.

The monomer fused to p3 and randomized on the DNA-binding interface was incorporated in the phage coat and the soluble I-CreI monomers produced in the same compartment either dimerize or interact with the displayed monomer, thereby producing particles displaying I-CreI homodimers (or heterodimers if the monomer fused to p3 was mutated).

Target production

Two complementary primers encoding the desired sequences but harboring an extra adenosine in 3' were annealed and ligated into pGEM-t Easy (Promega). After sequencing, a correct clone was chosen as template to PCR amplify a biotinylated 200 pb fragment using the kit KOD (Novagen) and primers SP6 (TTTAGGTGACACTATAGAATAC) and biotT7 (biot-TAATACGACTCACTATAGG). The PCR product concentration was estimated on gel and the fragment was used as such in ELISA or selection procedures.

Rescue of the phagemid library

A representative aliquot of the library (at least 10x more bacteria than the library size) was used to inoculate 50 ml of 2xTY containing 100 µg/ml ampicillin and 2% glucose (2TYAG) and the culture was agitated at 37°C. At an OD_{600} of 0.5, 5 ml of this culture was infected with helper phage K07 at a ratio phage:bacteria of 20:1 and incubated without agitation for 30 min at 37°C. After

centrifugation at 4000 rpm for 10 min at room temperature (RT), the pellet was resuspended in 25 ml of 2xTY containing 100 µg/ml ampicillin and 25 µg/ml kanamycin (2TYAK) and agitated overnight at 30°C. The culture was centrifuged at 4000 rpm for 20 min at 4°C and phage particles were precipitated by the addition of 0.2 volume of 20% PEG6000 /2.5M NaCl for 1h on ice.

After centrifugation at 4000 rpm for 20 min at 4°C, the phage pellet was resuspended in 1 ml of PBS and centrifuged at 10 00 rpm for 5 min. 0.2 volume of 20% PEG6000 /2.5M NaCl was added to the supernatant and the mix was centrifuged at 10 000 rpm to pellet the phage particles. Particles were finally resuspended in 250 µl PBS.

Selection procedure

Phage particles were diluted in 1 ml of PBS containing 3% dry milk and 25 mM CaC1 (PMC) and incubated for 1h at RT. 100 pI Streptavidin beads (Dynal, 200 µl for the first round) were washed 3X in PMC and blocked for 1h in the same buffer. The biotinylated targets were added to the phage at the indicated concentration and the mix was agitated at RT for 1h. Beads were added to the mix and incubated at RT for 15 min. Beads were collected on the vial wall using a magnet and washed 10X in PMC containing 0.1% tween. After a final wash in PBS, beads were resuspended in 0.5 ml of 100 mM Triethanolamine pH 12 and incubated for exactly 10 min. The supernatant were collected and immediately neutralized by 0.5 ml of 1 M Tris pH8. An aliquot of this eluate was serially diluted for titration and with 4 ml 2xTY. 5 ml of exponentially growing TG1 cells were added and the mix was incubated for 30 min at 37°C without agitation. Cells were plated on large 2TYAG plates and incubated overnight at 30°C. Colonies were resuspended in 2TYAG, adjusted to an OD_{600} of 100 and kept at —80°C after addition of 15 % glycerol.

Screening by phage ELISA

Isolated colonies from selection outputs were toothpicked into 100 µl of 2TYAG in 96 well plates, and agitated overnight at 37°C. Next day, a fresh plate containing 100 µl 2TYAG was isolated using a transfer device. 50 µl of sterile 60% glycerol was added to the overnight plate and this masterplate was stored at —80°C. The fresh plate was agitated at 37°C for 2.5 h, rescued by the addition of 2TYAG

containing 2x 10^9 pfu of helper phage M13K07, incubated for 30 min at 30°C, spun at 1700 rpm for 15 min. Cells pellets were resuspended in 150 µl 2TYAK and agitated overnight at 30°C. After centrifugation, 20 µl of supernatant was used as described in the previous section.

- Results

Selections

Phage particles displaying I-Cre I variants were produced by infecting bacteria harboring the phagemid library with helper phage M13KO7. Phage particles were purified by PEG precipitation and incubated with a biotinylated PCR fragment harboring HIV6335 target. After 1h of incubation at room temperature, streptavidin-coated magnetic beads were added to the solution to retrieve the biotinylated DMA and bound phages. The beads were extensively washed and the bound phages were eluted by pH shock. Bacteria were infected with the eluted phages and plated on large 2xTYplates containing ampicillin and 2% glucose. Serial dilutions of an aliquot of the eluted phages were used to infect bacteria to calculate the number of phage particle and obtain isolated colonies.

The day after, bacteria were scrapped from the large plates and stored as glycerol stocks. An aliquot (representative of the diversity) was used to produce a new batch of phage particles for a second round of selection.

The stringency of the selections was increased after each round. The first selection was done using 10 nM of biotinylated target. The second was done with 400 pM and the washing steps were extended. The third round was done using 250 pM and washed more extensively.

As shown on Table 1, the first and second rounds of selection against the HIV2 target lead to an output titer characteristic of background values (10^5 to 10^6 pfu/ml). However, a significant enrichment was measured on round 3.

Selection Round	Input (pfu/ml)	Output (pfu/ml)	Enrichment
1	$6.4x10^{11}$	$1.4x10^{5}$	NA
2	$4.0x10^{12}$	$3.0x10^{6}$	3
3	$2.8x10^{12}$	$6.9x10^{7}$	33

Table 1: Selection titers. C2H6335: selection done on HIV2 target using the library described in the other example. NA: non applicable. Enrichment is defined as (output n+1/input n+1)/(output n/input n).

Screening by phage ELISA

80 clones randomly picked from each output (as well as unselected clones) were used to produce phage particles displaying I-CreI variants in a mono-clonal fashion. Supernatants containing the phage particles were incubated on biotinylated PCR fragment immobilized on plastic via streptavidin. Bound phages were stained with an HRP-labeled anti p8 (major coat protein) monoclonal antibody (Pharmacia). As shown on Table 2, no binders were detected among the unselected clones or from the outputs of the first round of selection. However 60% of clones picked after round 2 against are positive against H6335 but negative on an irrelevant target (P1234, target of homing endonuclease PI-SceI). This result is in good agreement with the output titer. Indeed this selection only resulted in a mild enrich-ment, suggesting that a large number of clones still originate from background. As expected, a third round of selection lead to 99% of strong binders, which explains the large number of output phages after this third selection.

Selection round	% positive against C1234	% positive against P1234
0	0	0
1	0	0
2	60	0
3	99	0

Table 2 : Percentage of positive clones in a ELISA assay directed against the I-CreI target (C1234) or the HIV2 derived target (H6335). 77 clones were assayed for each output. Round 0: unselected library

Using phage display, new meganucleases were selected from a large library of I-Cre I variants. Selections on biotynilated DNA targets lead to an increase of output titers characteristic of an enrichment for molecules capable of binding the DNA targets. This enrichment was confirmed by phage ELISA. These results demonstrate the efficiency of the selection and screening methods.

A selection/screen experiment in yeast to identify novel meganucleases.

Material and methods

Bacterial and yeast strains

Every subcloning and plasmid preparations are performed in XLI-blue : *E.coli* provided by Stratagene following standard procedures. Experiments in *S. cerevisiae* are done in the following strains:

- FYC2-6A: alpha, trp1Δ63, leu2Δ1, his3Δ 200

- FYBL2-7B: a, ura3 Δ 851, trp1Δ63, leu2Δ1, lys2Δ202

- YASP3 (derived from FYC2-6A): alpha, ura3::SSA-ura3-HIV2-KanR, ade2::SSA-ade2-HIV2-TRP1, trp1Δ63, leu2Δ, his3Δ 200

Plasmids

- pCLS0279: ADH1 promoter, TRP1 selectable marker and ARS-CEN origin of replication, β-galactosidase SSA target, HIV2 6335 cleavage site.

- pCLS0569: kanamycin resistance cassette, HIV2 6335, internal fragment of the URA3gene.

- pCLS0570: kanamycin resistance cassette, HIV2 6335, internal fragment of the LYS2 gene.

- pCLS0576: TRP1 selectable marker, HIV2 6335, internal fragment of the ADE2 gene.

- pCLS0047: Galactose inducible promoter, LEU2 selectable marker and 2 micron origin of replication.

Results

An *in vivo* assay in yeast that allows to screen mutagenized I-CreI protein variants with detectable activity towards a specified target was performed.

A library of mutated I-CreI meganucleases has been first selected by a phage display procedure, resulting in a sub-library enriched for variants of interest, able to bind the HIV2 6335 target. The inserts from this enriched sub-library are subcloned into pCLS0047 under the control of a galactose-inducible promoter, for further selection in yeast. However, we can produce the library directly in the suitable yeast expression vector, and void the phage display step.

A specific yeast strain (YASP3) containing two reporter systems integrated in chromosomes was prepared. These two reporter systems are based on recombination by Single Strand Annealing (SSA). SSA is induced by specific cleavage of the HIV2 6335 site.

Namely, a URA3 SSA target and an ADE2 SSA target were introduced. The URA3 SSA target was a modified ura3 gene with 2 direct repeats of 600 base pairs separated by 4,3kb (containing a kanamycin resistance cassette and the HIV2 6335 cleavage site). The strain was unable to grow on a minimal medium lacking uracile but was resistant to G418. When this target was cleaved and recombined properly, the yeast was able to grow on media without uracil and was sensitive to G418.

The ADE2 SSA target was a modified ade2 gene with 2 direct repeats of 1,1kb separated by 3,6kb (containing a tryptophan selectable marker and the HIV2 6335 cleavage site). Because of this mutated ade2 gene, the yeast strain was unable to grow on a minimal medium lacking adenine, but harbored a red color on a medium with a low adenine content. Because of the tryptophan selectable marker, it was able to grow on minimal media without tryptophan. When this target was cleaved and recombined properly, the yeast was white, able to grow on media without adenine and unable to grow on a minimal medium lacking tryptophan.

Basically, the recipient yeast strain was red (on low adenine medium), G418 resistant, tryptophan prototroph and auxotroph for uracile and adenine. If a specific meganuclease is expressed in this strain and cleaves its target sites, the resulting yeast clone is white, G418 sensitive, prototroph for tryptophan and auxotroph for uracile and adenine.

The YASP3 strain was validated by determining the level of spontaneous recombination of each target alone and of both targets taken together. The

URA3 SSA 10 target recombined spontaneously as an uracile prototrophe, G418 sensitive at an approximate 6 10^{-4} rate. The ADE2 SSA target recombined spontaneously as an adenine prototrophe at an approximate 2,7 10^{-3} rate. Recombination of both markers occurred spontaneously (resulting in uracile/adenin rototrophes) at an approximate 10^{-6} rate.

A pilot experiment with 1,5x10 in transformants showed no background level of uracileladenine prototrophes means that the number of false positive clones should be less than 10 after a transformation experiment with a library that would yield about a million of independent clones.

The library is used to transform YASP3. A classical chemical/heat chock protocol that routinely gives 10^6 independent transformants per pg of DNA was used (Gietz, R.D. and Woods, R.A., 2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method (Methods Enzymol, 350, 87-96).

Transformation of the strain with the library gives more than 10^6 independent yeast transformants from which a number of clones are able to grow on a selective medium whithout uracile, leucine and containing galactose as a carbone source and a low amount of adenine. Among those clones, the interesting ones are white indicating that they contain a LEU2 vector allowing the expression of a meganuclease specific for HIV2 6335 site and that the enzyme is able to cut both URA3 and ADE2 reporters.

The positive clones are isolated and screened for their ability to induce the specific recombination of a plasmidic SSA β-galactosidase target (pCLSO279). This plasmidic reporter was a modified LacZ gene with 2 direct repeats of 825 base pairs separated by 1,3kb (containing a URA3 selectable marker and the HIV2 6335 cleavage site). The vector (which can be selected on a medium without tryptophan) is used to transform a yeast strain (FYBL2-7B) and clones are maintained on minimal media 35 lacking uracile to maintain the unrecombined LacZ target.

Yeast clones resulting from the selection experiment are mated with the yeast strain containing the SSA β-galactosidase target. Diploids are selected and

assayed for induced β-galactosidase activity. A number of clones are expected to behave as false positives at this step. They correspond to the background level of spontaneous recombination of the URA3 and ADE2 SSA targets. All remaining clones (uracile and adenine auxotrophes able to induce recombination of the SSA-LacZ target) are true positives expressing a meganuclease cleaving the HIV2 6335 target in vivo. Also, other experiments, based on the ones described above, can be used to determine more precisely the activity of such novel enzymes.

Example 3: Use of meganuclease for antiviral therapy

- Experimental Procedures

Cells

COS-7 cell lines from the american Type culture collection (ATCC) were cultured in DMEM plus 10% fetal bovine serum. PC-12 cells from ATCC were grown in RPMI1640 supplemented with 10% heat-inactivated horse serum and 5% heat-inactivated fetal bovine serum. PC-12 cells were differentiated as previously described (Su et al., 1999, Journal of Virology, 4171-4180). Briefly, cells were seeded on 6 well-plate at 5 10^4 cells per well. The following day, cells were incubated in PC-12 medium containing 100 ng/ml of 2,5S NGF (Invitrogen). Medium was changed every three days. After 7 days of incubation, undifferentiated cells were eliminated by adding 2μM of fluorodeoxyuridine (FdUrd).

Construction of recombinant HSV-1

HSV-1 was purchased from ATCC. Viruses were propagated on COS-7 cells at low MOI (0,01PFU/cell). Recombinant virus (rHSV-1) were generated as previously described (Lachmann, R.H., Efstathiou, S., 1997, Journal of Virology, 3197-3207). A 4,6 Kb pstI-bamHI viral genomic DNA fragment was cloned in pUC 19. Based on HSV-1 sequence from data base (ID : NC 001806), this region represents nucleotides 118867 to 123460. A cassette consisting of a CMV promoter driving Lac gene expression was introduced into a168 bp HpaI deletion. This region is located within the major LAT locus of HSV-1. I-Sce I cleavage site was finally cloned directly after the CMV promoter. This construct was used to generate recombinant viruses. Plasmid was linearized by XmnI digestion, and 2μg of this plasmid DNA was contransfected with 15μg of HSV-1 genomic DNA prepared from COS-7 infected cells by $CaCl_2$ method. After 3 or 4 days, infected cells were harvested and sonicated.

Aliquot of the lysed cells were used to infect COS monolayer. Virus recombinant were selected by overlaying COS monolayer with 1% agarose in medium containing 300µg/ml of X-Gal (5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside). Blue clones were picked and further subjected to three round of plaque purification. Presence of the I-Sce I site was confirmed by PCR and in vitro I Sce-I enzymatic digestion.

Viral inhibition

6 well-plate were seeded with 2.10^4 cells per well. The next day COS-7 cells were transfected with 0,5 µg of plasmid expressing ISce-I or containing the ISce-I ORF in the opposite orientation by the EFFECTENE method according to the manufacturer protocol. We achieved routinely in our laboratory 60 to 70% efficiency using this methodology. Fourthy eight hours later, subconfluent transfected cells were infected with rHSV-1. For infection, rHSV-1 was diluted in PBS containing 1% fetal bovine serum and adsorbed onto cells for 20-40 min at 37°, in humidified incubator with 5% CO2. 6 wells-plates were infected at 30 or 300 PFU per well for respectively viral inhibition or cells survival experiments. Cells were harvested at day 1, 2, and 3 and ß-galactosidase activity was assayed and DNA extracted.

ß-galactosidase activity

Cell monolayer was fixed in 0,5% glutaraldehyde in 100mM PBS containing 1mM $MgCl_2$ at 4° for 10 minutes. After one wash with detergent solution (100mM PBS, 1mM $MgCl_2$, 0,02% Nonidet p-40) cells were incubated at 37° in X-Gal stain solution (10 mM PBS, 1mM $MgCl_2$, 150 mM NaCl, 33 mM $K_4Fe(CN)_6.3H_2O$, 33 mM $K_3Fe(CN)_6$, 0,1% X-Gal) until color development. Beta-galactosidase activity was also measured on cell extract with o-nitrophenyl-ß-D-galactopyrannoside (ONPG) as substrate. Cell monolayer was washed once with PBS. Cells were then lysed with 10 mM Tris ph 7.5, 150 mM NaCl, 1% Triton X-100, protease inhibitors. After 30 minutes incubation on ice cell lysate was centrifuged and ß-galactosidase was assayed. Typically 30µl of supernatant was combined with 270 µl of reaction buffer (10 mM PBS; ph 7,5, 1 mM $MgCl_2$, 0,3% ß-mercaptoethanol) containing 800µg /ml ONPG. The reaction was carried out at 37° and stopped with 0.5 ml of 1M $NaCO_3$. Optical density was measured at 415 nm. Beta-galactosidase activity is calculated as relative unit normalized for protein concentration and incubation time.

Semi-quantitative PCR

To measure viral replication of rHSV-1, oligonucleotides were designed to amplify a 217 bp fragment from Lac gene. The standard DNA used in this assay was generated by cloning this fragment in a Bluescript plasmid, and by inserting a 50 bp fragment downstream to the 5' oligonucleotide. PCR of the standard produced 267 bp amplicon. Series of PCR (not shown) were carried out to fix the amount of standard and DNA sample, and the number of cycles to achieve linear response of the amplification. The basic semi-quantitative PCR were carried out in a total volume of 30µl, using the READYMIX™ TAQ (Sigma) with 20 pmols of each primers and 180 pg of DNA. The tubes were heated for 4 min at 94° and subjected to 22 cycles : 94° for 1 min, 62° for 50 sec, 72° for 2 min, and 72° for 7 min.

Virus titration

In one series of experiments, the culture medium was collected every day at days 1, 2,3, and 4, and fresh medium was added. In the other, the medium was not changed during experiment and aliquots were collected every day. To monitor for the release of HSV-1 progeny, aliquot of medium were titred on COS-7 cells by standard plaque assay.

- Results

The effect of I-Sce I on viral replication was examined using a recombinant Herpes simplex virus carrying a I-Sce I restriction site (rHSV-1). For convenience, rHSV-1 was build with a cassette containing CMV promoter driving the Lac gene. I-Sce I site was inserted at the junction of the CMV promoter and Lac gene. The expression cassette was cloned by homologous recombination in the major LAT locus which allowed Beta-galactosidase (ß-gal) expression during lytic infection in COS-7 cell monolayer. Strinkingly transfection of I-Sce I expression vector before viral infection virtually completely inhibited HSV-1 plaque formation in COS cells (Figure 5) as shown by X-Gal coloration. In contrast, control transfection with expression vector containing I-Sce I open reading frame in the reverse orientation which did not allow any fonctional transcript, did not affect viral replication. Furthermore, 48 hours after infection, the cells were checked for I-Sce I expression. All the lysis plaques formed in cells monolayer transfected with I-Sce I expression vector represented cells which did not expressed I-Sce I (the transient transfection is about 70%

efficient). However, cells expressing I-Sce I surrounding the lysis plaque inhibited the viral propagation. The I-Sce I effect was confirmed by measuring the ß-galactosidase activity in a cell lysate. After infection of COS-7 cells monolayer transiently expressing I-Sce I with 30 Pfu per well cell monolayer was collected at day 1, 2, and 3 post infection and ß-gal was assayed. Figure 6A shows a drastic decrease of the ß-galactosidase activity reflecting the inhibition of rHSV-1 replication. The protective effect of I-Sce I over a time course of rHSV-1 infection was evaluated next. At 3 days after infection, cells transfected with I-Sce I expressing vector shown no sign of cytopathic effect whereas control cultures were completely lysed as shown in Figure 6B. In an effort to quantify the degree of inhibition of viral DNA replication by I-Sce I, we have set-up a semi-quantitative PCR. Genomic DNA was extracted from cells at day 1, 2, and 3 after infection. PCR was carried out with primers a and b (Figure 7A) generating a 217 bp amplicon in Lac gene. An internal standard was added in sample before PCR to quantify DNA. Lac gene was virtually not detectable in I-Sce I expressing cells at 3 days post-infection (Figure 7A). In contrast cells that did not received I-Sce I expression vector shown high levels of virus DNA. This result was confirmed by PCR using primers in viral endogenous gene (Figure 7B). Amplification of Thymidine Kinase (TK) gene shown that I-Sce I inhibited the viral replication. Finally COS–7 cells expressing active I-Sce I or I-Sce I ORF in the reverse orientation were infected with rHSV-1 and the concentration of virus released in the medium at different time points was measured by plaque assay (figure 8). Viruses were quantified in a rough array at day one when I-Sce I was produced. Viruses production was still markly decreased two days after the infection when compared with cells which did not expressed I-Sce I showing that I-Sce I effectively inhibited viral replication. This effect was still observed at day three although in a lesser extent. Probably the high mutation rate occuring during viral replication allowed emergence of mutant HSV-1 which were able to escape the I-Sce I activity.

Taking together, these results demonstrates that I-Sce I and more generally meganucleases can be used to inhibit viral infection. The use of custom-made meganuclease or combination of custom-made meganucleases designed to cut specific viral sequences could represent a powerfull new strategy in the antiviral therapy.

Example 4: Meganuclease with altered binding properties derived from I-CreI homing endonuclease

The purpose of this experiment was to obtain novel meganucleases binding target sites close to the I-*Cre*I natural target site. A series of 6 targets were used (Figure 9), including the wild-type natural I-*Cre*I target (named C1234), the HIV2 target described in example 2 (named here H1234), and four additional targets. These four additional targets are 24 bp palindromes corresponding to inverted repeats of a 12 bp half I-CreI or HIV2 target site: C1221 and C4334 are inverted repeats of the first half and second half, respectively, of the C1234 target; H1221 and H4334 are inverted repeats of the first half and second half, respectively, of the H1234 target.

In contrast with example 2, the method used here did not involve any selection step, but was based on the extensive screening of the Lib2 library (see example 2). Three residues (Q44, R68 and R70) capable of base specific interactions with the DNA target were selected. The combinatorial library was obtained by replacing the three corresponding codons with a unique degenerated VVK codon. Eventually, mutants in the protein library corresponded to independant combinations of any of the 12 amino acids encoded by the VVK codon (ADEGHKNPQRST) at three residue positions. In consequence, the maximal (theoretical) diversity of the protein library was 12^3 or 1728.

Materials and Methods

Construction of a phage-displayed library of I-CreI variants.

First, residue D75, which is shielded from solvent by R68 and R70, was mutated to N (Asn) in order to remove the likely energetic strain caused by replacements of those two basic residues in the library. Homodimers of mutant D75N (purified from *E. coli* cells wherein it was over-expressed using a pET expression vector) were shown to cleave the I-*Cre*I homing site. A phagemid vector was then engineered that encodes the D75N mutant (Figure 2C : « Template ») fused to the phage coat protein p3 and phage-displayed D75N monomers were shown to bind the I-CreI natural DNA target (C1234 on Figure 9).

Then, DNA fragments carrying combinations of the desired mutations were obtained by PCR (several reactions in 50μl), using degenerated primers (Figure 2D : UlibIIfor, UlibIIrev) and as DNA template, the D75N gene. Lib2

was constructed by ligation of the corresponding PCR products, digested with specific restriction enzymes, into the D75N mutant gene, within the phagemid vector, as described in example 2.

Screening of meganucleases binding to the 6 different targets

Screening was performed by Phage ELISA, as described in example 2.

- Results

4560 clones (more than 2.5 times the theoretical mutant library diversity) were individually picked and screened by phage ELISA with the 6 different targets. 28 positives (clones binding one of the six targets) were identified. For validation, these 28 clones were re-assayed by phage ELISA, 8 times in parallel with the 6 different targets; 20 clones were thus confirmed as true positives. Finally, all 28 clones were sequenced.

Class I	Class II	Class III	Class IV
Q R K	(NTQH)N	Q R T (2)	Unknown sequence
Q R R	Q R N	(RG) (ED)	
H (KEQ) E	Q R A	Q Q K (2)	
	Q S R	Q N K	
	Q T R (2)		
	Q Q R		
	Q H K		
	D S H		
	Unknown sequence		

Table 3: Sequence of the proteins found in the four different classes. Only amino acids from position 44, 68 and 70 are indicated. Clones found twice are labeled with (2).

Four different patterns (ELISA results) could be observed. Figure 10 features one representative example for each one. The first class (Class I) corresponds to a strong binding of C1234, C1221, C4334 and H4334. The wild-type protein (QRR) was recovered in this class, showing that Class I profile is the regular binding profile of the original scaffold. Two variants were also shown to display such binding (QRK and another yet not completely identified mutant).

Variants from the second class have lowered their affinity for all targets, but H4334, since no binding was observed with C1234, C1221 and C4334. Eight different proteins were found to belong to this class, plus a protein which sequence could not be determined. Among the sequence variants of Class II, five retain the Q44 amino acid from the wild-type sequence, and one of the two arginines in position 68 or 70. However, in one mutant (DSH), none of the amino acids from position 44, 68 and 70 has been retained. Class III (4 different proteins) has a more complex pattern, as it retains apparent binding for the C1221 and H4334 target. Finally, one protein (Class IV) retains only a slight binding for target C1221 as none of the other targets are bound anymore.

It is difficult to draw conclusions from Class IV, since the residual binding with C1221 is very low, and sequencing of the unique Class IV mutant has failed. However, comparison of Class II and III with the wild-type profile of Class I clearly shows that the binding specificity has been altered.

The conclusion is that even from small libraries such as Lib2 (complexity 1.7 10^3), variants with altered binding profiles can be isolated, as shown in figure 10. Therefore, strategies based on screening, starting with larger mutant libraries, should allow the identification of more dramatic alterations, for instance binding for targets that were not bound by the initial protein scaffold. In addition, this approach leads to the identification of many different proteins for each profile. An extensive study of this kind should also bring the basis of a better understanding of DNA/meganuclease interactions.

Exemple 5: Comparison of selection and screening methods in yeast library

The purpose is here to compare the screening and selection methods in yeast. Whereas screening is the extensive examination of each individual clone of a population for its desired properties (for us, the cleavage properties), selection is an enrichment step : an initial library is submitted to the selection process, resulting in a sublibrary enriched for clones with the desired properties.

Since the throughput of the screening process can be insufficient to process very large number of clones, one or several selection steps can be useful when one has to deal with a very high diversity. However, selection can bring several unexpected bias, resulting in the selection of other properties than the ones wished by

the operator. Thus, it is extremely important to validate any selection process carefully.

Therefore, a selection method and a screening method were developped to look for meganuclease cleaving specific DNA target in yeast. Both are based on the production, by homologous recombination, and more precisely, by Single-Strand annealing, of specific markers, upon cleavage of the DNA target by the meganuclease within the yeast cell. The principle of these assays is described in Example 2. Selection is based on the restoration of an auxotrophy marker, (*URA3* in Example 2, *ADE2* and *LYS2* in this example), whereas screening is based on the restoration of a color marker, (LacZ and *ADE2* in example 2, LacZ in this example (Since an *ade2* mutation results in no growth, or in a yeast red color, depending on the amount of adenine in the culture medium, it can be used for both selection and screening).

Thus, the clones screened as positive with and without selection on a small library were compared, in order to check whether the selection method is suitable.

- Material and methods

Bacterial and yeast strains

Every subcloning and plasmid preparations are performed in XL1-blue: E.coli provided by Stratagene following standard procedures.

Experiments in *S. cerevisiae* are done in the following strains, wherein *cutI-CreI* represents the cleavage site for I-CreI:

- FYC2-6A: *alpha, trp1Δ63, leu2Δ1, his3Δ200*
- FYBL2-7B: *a, ura3Δ851, trp1Δ63, leu2Δ1, lys2Δ202*
- YAP4 and YDD6 (derived from FYC2-6A): *alpha, lys2::SSA-ura3-cutI-CreI-KanR, ade2::SSA-ade2-cutI-CreI-TRP1, trp1Δ63, leu2Δ1, his3Δ200*

Plasmids

- pCLS050: *ADH1* promoteur, *TRP1* selectable marker and ARS-CEN origin of replication, β-galactosidase SSA target, I-CreI cleavage site.
- pCLS0047: Galactose inducible promoter, *LEU2* selectable marker and 2 micron origin of replication.

- Results

A library of mutated I-*Cre*I meganucleases, namely Lib2 (see example 4) was introduced into a yeast vector (pCLS0047).

Two specific yeast strain (YAP4 and YDD6) containing two reporter systems integrated in chromosomes were prepared. These two reporter systems are based on recombination by Single Strand Annealing (SSA). SSA is induced by specific cleavage of the I-*Cre*I site.

Namely, a *LYS2* SSA target and an ADE2 SSA target were introduced. The *LYS2* SSA target was a modified *lys2* gene with 2 direct repeats of 3240 bp base pairs separated by 4,3kb (containing a kanamycin resistance cassette and the I-*Cre*I cleavage site). The strain was unable to grow on a minimal medium lacking lysine but was resistant to G418. When this target was cleaved upon overexpression of I-*Cre*I. and recombined properly, the yeast was able to grow on media without lysine and was sensitive to G418.

The ADE2 SSA target was a modified *ade2* gene with 2 direct repeats of 1,1kb separated by 3,6kb (containing a tryptophan selectable marker and the I-*Cre*I cleavage site). Because of this mutated *ade2* gene, the yeast strain was unable to grow on a minimal medium lacking adenine. Because of the tryptophan selectable marker, it was able to grow on minimal media without tryptophan. When this target was cleaved and recombined properly, the yeast was able to grow on media without adenine and unable to grow on a minimal medium lacking tryptophan.

Basically, the recipient yeast was G418 resistant, tryptophan prototroph and auxotroph for both lysine and adenine. If a specific meganuclease is expressed in this strain and cleaves the I-*Cre*I target sites, the resulting yeast clones are G418 sensitive, auxotroph for tryptophan and prototroph for lysine and adenine.

The Lib2 library was used to transform YAP4 and YDD6. Aclassical chemical/heat chock protocol that routinely gives us 10^6 independent transformants per µg of DNA was used(Gietz and Woods, 2002, Methods Enzymol, 350, 87-96).

Transformation of the strain with the library gives more than 10^6 independent yeast transformants on a selective medium whithout leucine (selection for the meganuclease expression vectors), and containing glucose as a carbone source and a low amount of adenine. Such clones can be screened after mating with the yeast

strain containing the SSA β-galactosidase target (FYBL2-7B transformed with pCLS050). Diploids are selected and assayed for induced β-galactosidase activity. Screening of 2400 independant clones resulted in the identification of 15 positives. DNA from these clones was recovered, electroporated into *E. coli*, for recovery of the meganuclease expression plasmid. For each of these 15 positives, *two E. coli* clones were amplified and sequenced. No difference of sequence was observed between two *E. coli* clones obtained from the same positive yeast clone. Plasmids were then retransformed into the yeast YAP4 or YDD6 strain, and screening was done again. It confirmed the results of the primary screen for the 15 different yeast clones.

The same screening experiment was then performed, with the exception that a selection step was added. The Lib2 libary was transformed into YDD6, and the cells were plated onto a a selective medium without leucine (selection for the meganuclease expression vectors), and containing glucose as a carbone source. After three days of growth, colonies were resuspended in water, and plated onto a a selective medium whithout leucine (selection for the meganuclease expression vectors), adenine and lysine (selection for the clones wherein the two ADE2 and LYS2 SSA reporter had recombined) and containing galactose as a carbone source. 960 clones were obtained, and 845 (88%) were screened as positive after mating with the yeast strain containing the SSA β-galactosidase target (FYBL2-7B transformed with pCLS050). Eighteen out of them were confirmed by a second round of screening.

These results gave a measure of the enrichment in a single round of selection. Since screening gave us 15 positives out of 2400 clones without selection and 845 out of 960 with selection, the enrichement is (845/960)/(15/2400)=140.

Then the 33 confirmed clones (18 obtained with and 15 obtained without selection) were sequenced. The table below describes the clones obtained by selection and screening. Since Lib2 results from mutation of residues Q44, R68 and R70 of I-*Cre*I, clones are described by three letters, corresponding to the amino acids (one letter code) presents at positions 44, 68, and 70. For example, wild type would appear as QRR, as a TRR mutant corresponds to a single mutation at position 44, replacing the glutamine with a threonin.

Mutant I-CreI protein	Number of clones containing the mutation after simple screening of 2500 clones	Number of clones containing the mutation after selection AND screening
TRR	6 clones (out of 15)	12 clones (out of 18)
QRA	3 clones (out of 15)	5 clones (out of 18)
QAR	2 clones (out of 15)	
TRK	1 clone (out of 15)	
TRA	1 clone (out of 15)	
TRN	1 clone (out of 15)	1 clone (out of 18)
ARN	1 clone (out of 15)	

Table 4 : Comparison of the clones obtained either by simple screening or selection and screening.

Clearly, the selection process is very efficient, since an enrichement factor of 140 is observed after selection. Among the positives, clones TRR and QRA are the most frequent with and without selection. Four sequences were found after simple screening, and not after selection. However, these differences are not statistically significant: only 18 of the 845 positives found after selection were sequenced and haracterization of a larger sample should identify more different positive variants.

These results show that with a small library such as Lib2, positives can be selected efficiently, whereas the biases that could be introduced by this selection system are small if any. Thus, this selection procedure can be used to enrich larger libraries for variants of interest, for example when these libraries cannot be entirely processed by the screening method. In this example, the mutants that have still the ability to cleave the I-*Cre*I target site were recovered. However, mutants cleaving novel targets could be obtained just in the same way.

Example 6: Screening of active meganuclease based on their clevage properties *in vitro*, and in mammalian cells.

The purpose of this experiment is to demonstrate that active and inactive meganucleases can be distinguished on simple screening assays *in vitro* and in mammalian cells. The *in vitro* assay is similar to a restriction assay, and the assay in

mammalian cells is based on cleavage-induced recombination. Cleavage in mammalian cells is similar to the assay in yeast (example 2): cleavage induced recombination (and more precisely single-strand annealing) results in a functional LacZ reporter gene which can be monitored by standard methods. However, in contrats with the yeast assay, which relies on stable replicative plasmids, the cell-based assay is working with transient matrix.

As for examples 4 and 5, the Lib2 library was used. Three residues (Q44, R68 and R70) capable of base specific interactions with the DNA target were selected. The combinatorial library was obtained by replacing the three corresponding codons with a unique degenerated VVK codon. Eventually, mutants in the protein library corresponded to independant combinations of any of the 12 amino acids encoded by the VVK codon (ADEGHKNPQRST) at three residue positions. In consequence, the maximal (theoretical) diversity of the protein library was 12^3 or 1728. The target used was the natural I-*Cre*I target (named C1234).

2000 clones were individually picked, and tested with the two screening methods, to establish a comparizon of these methods. Positives were then sequenced and analyzed.

- Materials and Methods

Construction of a library of I-CreI variants for expression in vitro.

First, residue D75, which is shielded from solvent by R68 and R70, was mutated to N (Asn) in order to remove the likely energetic strain caused by replacements of those two basic residues in the library. Homodimers of mutant D75N were shown to cleave the I-*Cre*I homing site. The library was constructed in the pTriex vector (Novagen).

Then, DNA fragments carrying combinations of the desired mutations were obtained by PCR (several reactions in 50 µl), using degenerated primers (Figure 2D : UlibIIfor, UlibIIrev) and as DNA template, the D75N gene. Lib2 was constructed by ligation of the corresponding PCR products, digested with specific restriction enzymes, into the D75N mutant gene, within the pTriex vector, as described in example 2.

Production of meganucleases in vitro

50 ng of DNA plasmid solution in 2 µl were mixed to 4 µl of RTS solution (Rapid Translation System RTS 100, E. coli HY kit from Roche). Protein productions were done *in vitro* at 30°C for at least 4 h. After production of the proteins, the fresh preparation was diluted in distillated water prior to try the meganuclease activity.

In vitro Cleavage assay

pGEMT plasmids with single meganuclease DNA target cut sites were first linearized with *Xmn*I. Cleavage assays were performed at 37°C in 12.5 mM HEPES (pH 8), 2.5% (v/v) glycerol and 10 mM MgCl2. Reactions were stopped after 1 hour by addition of 0.1 volume of 0.1 M Tris–HCl (pH 7.5), 0.25 M EDTA, 5% (w/v) SDS, and 0.5 mg/ml proteinase K and incubation at 37°C for 20 minutes. Reaction products were examined following separation by electrophoresis in 1% agarose gels.

Cleavage in mammalian cells

CHO cells were cotransfected with the meganuclease expressing pTriEx plamid and the reporter plasmid. The reporter plasmid contains an inactive LacZ gene under the contol of an appropriate promoter. LacZ is inactive because it contains an internal duplication of 220 bp, and an insertion of the target site (24 to 80 pb), located between the two 220 bp repeats. For transfection, Polyfect transfection reagent was used accordingly to the supplier (Qiagen) protocol. 72 hours after transfection, tissue culture medium was removed and cells were incubated in lysis buffer (Tris-HCl 10 mM pH7.5, NaCl 150 mM, Triton X100 0,1%, BSA 0.1 mg/ml, protease inhibitors). The whole lysate was combined with 0,1 volume of Mg 100X buffer ($MgCl_2$ 100 mM, β-mercaptoethanol 35%), 1,1 volume of ONPG 8 mg/ml and 7,8 volume of sodium phosphate 0.1 M pH7.5. After incubation at 37°C, the reaction was stopped with 0,5 volume of 1 M Na_2CO_3 and OD was measured at 415nm. The relative □ -galactosidase activity is determined as a function of this OD. Positives are clones choosen by comparizon with a negative control where an empty expression vector is tranfected to the cells. Their β-galactosidase activities is higher than (M + 2 E) (where M is the average β-galactosidase activity of the negative controls and E the standard deviation between those same mesurements).

- Results

2000 clones clones (more than the theoretical mutant library diversity) were individually picked and cultured in 96 deep-well plates. Plasmid DNA was extracted using a BioRobot8000 platform (Qiagen) with the Qiaprep 96 Turbo BioRobot kit (Qiagen). Since the pTriEx plamid has been designed to drive the expression of protein in bacteria and mammalian cells, plasmid DNA was then used for both the *in vitro* assay and the assay in cells. All the positives obtained with either methods were rechecked for cleavage *in vitro* and in cells.

For screening with the *in vitro* assay, Meganuclease were produced *in vitro* from each individual plasmid with the RTS (Roche) system, and then tested for cleavage of a linearized pGEMT (InvitroGene) plasmid containing the targets. The digests were then run on an electrophoresis gel to detect the expected cleavage products.

For screening with the assay in mammalian cells, the plasmid were cotransfected with reporter plasmid containing the target sites in CHO cells, and cleavage-induced recombination was monitored 72hours later, as a function of □ -galactosidase activity.

Out of the 2000 clones, 206 clones (10.3%) were found to be positive with the *in vitro* assay. In mammalian cells, 85 clones (4.3%) were found to be positives. Out of these 291 clones, 82 were positive in both methods.

121 different variant proteins were identified after sequencing of positive clones. The identity and cleavage properties of these 121 variants are shown in the table below. 54 variants display a detectable cleavage activity in both assays, as 66 are positives only for cleavage *in vitro*. A single variant is positive in the cell-based assay but not *in vitro*.

Residue 44	Residue 68	Residue 70	in vitro cleavage	SSA in cells
Ala	Arg	Ala	+	+
Ala	Arg	Arg	+	+
Ala	Arg	Gly	+	+
Ala	Arg	His	+	+
Ala	Arg	Ser	+	+
Ala	Lys	Lys	+	+
Ala	Thr	Lys	+	+
Asn	Arg	Arg	+	+
Asn	Arg	His	+	+

Asn	Arg	Pro	+	+
Asp	Pro	Thr	+	+
Gln	Ala	Arg	+	+
Gln	Ala	His	+	+
Gln	Arg	Ala	+	+
Gln	Arg	Arg	+	+
Gln	Arg	Asn	+	+
Gln	Arg	His	+	+
Gln	Arg	Pro	+	+
Gln	Arg	Ser	+	+
Gln	Arg	Thr	+	+
Gln	Asn	Arg	+	+
Gln	Gln	Arg	+	+
Gln	Gln	Thr	+	+
Gln	His	Arg	+	+
Gln	His	Asn	+	+
Gln	His	Gln	+	+
Gln	Lys	Ala	+	+
Gln	Lys	Gln	+	+
Gln	Lys	Lys	+	+
Gln	Lys	Thr	+	+
Gln	Ser	Gly	+	+
Gln	Ser	His	+	+
Gln	Thr	Arg	+	+
Glu	Arg	Arg	+	+
Pro	Arg	Gly	+	+
Pro	Arg	Thr	+	+
Ser	Arg	Asn	+	+
Ser	Arg	Lys	+	+
Ser	Lys	Arg	+	+
Ser	Thr	Arg	+	+
Thr	Ala	Arg	+	+
Thr	Ala	Thr	+	+
Thr	Arg	Ala	+	+
Thr	Arg	Arg	+	+
Thr	Arg	Asn	+	+
Thr	Arg	Gly	+	+
Thr	Arg	Lys	+	+
Thr	Arg	Ser	+	+
Thr	Arg	Thr	+	+
Thr	Gln	Arg	+	+
Thr	Gly	Arg	+	+
Thr	Lys	Arg	+	+
Thr	Thr	Arg	+	+
Thr	Thr	Lys	+	+
Ala	Ala	Arg	+	-
Ala	Arg	Asn	+	-
Ala	Arg	Thr	+	-

Ala	Gln	Arg	+	-
Ala	Gln	Lys	+	-
Ala	Glu	Asn	+	-
Ala	Gly	Arg	+	-
Ala	His	Arg	+	-
Ala	Ser	Lys	+	-
Ala	Thr	Arg	+	-
Arg	Ala	Ser	+	-
Arg	Arg	Asn	+	-
Arg	Thr	Asn	+	-
Asn	Ala	Arg	+	-
Asn	Arg	Asn	+	-
Asn	Arg	Gly	+	-
Asn	Arg	Ser	+	-
Asn	Arg	Thr	+	-
Asn	Asn	Arg	+	-
Asn	Gln	Arg	+	-
Asn	Thr	Arg	+	-
Asp	Arg	Lys	+	-
Asp	Ser	Asp	+	-
Gln	Ala	Ala	+	-
Gln	Ala	Ser	+	-
Gln	Ala	Thr	+	-
Gln	Asn	Asn	+	-
Gln	Asn	Gly	+	-
Gln	Asn	Pro	+	-
Gln	Asn	Thr	+	-
Gln	Gln	Asn	+	-
Gln	Lys	Asp	+	-
Gln	Pro	Asn	+	-
Gln	Ser	Ala	+	-
Gln	Ser	Asn	+	-
Gln	Thr	Ser	+	-
Gly	Ala	Arg	+	-
Gly	Arg	Ala	+	-
Gly	Arg	Arg	+	-
Gly	Arg	Gly	+	-
Gly	Arg	Ser	+	-
Gly	Arg	Thr	+	-
Gly	Gln	Arg	+	-
Gly	Gln	Lys	+	-
Gly	His	Arg	+	-
His	Ala	Arg	+	-
His	Arg	Ala	+	-
His	Arg	Thr	+	-
Lys	Ala	Ser	+	-
Lys	Ala	Thr	+	-
Lys	Arg	Gln	+	-

Lys	Arg	Thr	+	-
Pro	Asn	Thr	+	-
Ser	Arg	Ala	+	-
Ser	Arg	Gly	+	-
Ser	Arg	His	+	-
Ser	Arg	Thr	+	-
Ser	Asn	Arg	+	-
Ser	His	Arg	+	-
Ser	Ser	Arg	+	-
Ser	Ser	Lys	+	-
Thr	Arg	Asp	+	-
Thr	Arg	His	+	-
Thr	Gln	Lys	+	-
Thr	His	Arg	+	-
Thr	Ser	Arg	+	-
Gln	His	His	-	+

Table 5 : Sequence end cleavage properties of the different variants

These results clearly show the discrimination of positive and negative clones can be achieved in two different assays, based on cleavage *in vitro* and cleavage in mammalian cells. Cleavage *in vitro* seems to be more sensitive, since much more positives were identified. These differences can be due to many factors resulting either from meganuclease expression, or activity, either from the detection based on recombination. However, all but one clone (Gln44 His68 His70) identified in cells appeared to be confirmed *in vitro*, which shows that the cell-based assay is reliable to identify functional endonucleases.

Example 7 : Meganuclease-induced recombination of an extrachromosomal reporter *in toto* using I-Sce I expressing plasmid

A-Optimization of the reporter system

- Experimental procedures

Vectors construction

The target vectors are based on a *LagoZ* expression vector driven by promoter of the human EF1-alpha gene. This promoter has been shown previously to have wide expression spectrum *in vivo* (Kim, D.W., Uetsuki, T., Kaziro, Y., Yamaguchi, N., Sugano, S, 1990, Gene, 91, 217-223). The promoter region includes splice donor and acceptor sites in the 5' untranslated region of the h-EF1-alpha gene-*LagoZ* is a CpG island depleted *LacZ* gene designed to abolish gene silencing in transgenic mice (Henry I, Forlani S, Vaillant S, Muschler J, Choulika A, Nicolas JF,

1999, C R Acad Sci III. 322, 1061-70). To construct target vectors with different lengths of homology, the 3' fragment of the *LagoZ* gene was first deleted (about 2000 bp) and replaced by the *I-Sce* I cleavage site. The 3' fragments of different lengths were generated by digestion of the parental plasmid. These fragments contained different amounts of homology with the 5' fragment of the *LagoZ* gene. Finally these DNA fragments were individually cloned adjacent to the *I-SceI* cleavage site, creating different target vectors with 0, 70, 220, 570, and 1200 bp of homology, respectively.

Cell culture

COS-7 and CHO-K1 cell lines from the American Type Culture Collection (ATCC) were cultured in DMEM or Ham's F12K medium respectively plus 10% fetal bovine serum. For *I-Sce I* induced Single Strand annealing (SSA) assays, cells were seeded in 12 well-plates at a 15.10^3 cells per well one day prior transfection. Transient transfection was carried out the following day with 500 ng of DNA using the EFFECTENE transfection kit (Qiagen). Equimolar amounts of target plasmid and *I-SceI* expression vector were used. The next day, medium was replaced and cells were incubated for an other 72 hours.

ß-galactosidase activity

Cell monolayers were fixed in 0,5% glutaraldehyde in 100mM PBS containing 1mM $MgCl_2$ at 4° for 10 minutes. After one wash with detergent solution (100mM PBS, 1mM $MgCl_2$, 0,02% Nonidet p-40) cells were incubated at 37° in X-Gal stain solution (10 mM PBS, 1 mM $MgCl_2$, 150 mM NaCl, 33 mM $K_4Fe(CN)_6.3H_2O$, 33 mM $K_3Fe(CN)_6$, 0,1% X-Gal) until color development. Beta-galactosidase activity was also measured in cell extracts with o-nitrophenyl-ß-D-galactopyrannoside (ONPG) as a substrate. Cell monolayers were washed once with PBS. Cells were then lysed with 10 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X-100, protease inhibitors. After 30 minutes incubation on ice, cells lysates were centrifuged and ß-galactosidase was assayed. Typically 30 µl of supernatant was combined with 270 µl of reaction buffer (10 mM PBS; pH 7,5, 1 mM $MgCl_2$, 0,3% ß-mercaptoethanol) containing 800 µg /ml ONPG. The reaction was carried out at 37° and stopped with 0,5 ml of 1M $NaCO_3$. Optical density was measured at 415 nm. Beta-galactosidase activity is calculated as relative unit normalized for protein concentration and incubation time.

- Results

When a DNA double-strand break (DSB) is introduced between two repeated sequences, it induces homologous recombination resulting in a deletion of the repeats, together with all the intervening sequences. The recombination pathway is often referred to as the single-strand annealing (SSA) pathway. A reporter system was designed to monitor meganuclease-induced SSA in animal models *in toto*. In order to optimize the reporter system, the correlation between meganuclease-induced SSA efficiency and repeat length was first examined. Different target vectors carrying a *LagoZ* gene containing duplications of various sizes were constructed (Figure 11). The presence of the duplication and of the *I-SceI* cleavage site inactivates the gene. The repair of the *LagoZ* gene by SSA results in the loss of one repeat and of the cleavage site, and in the restoration of a functional *LagoZ* gene. *LagoZ* codes for the β-galactosidase enzyme which can be detected by colorimetry. Transient transfection with equimolar amounts of target vector and *I-SceI* expression vector or expression vector that doesn't express the meganuclease were carried out in CHO or COS-7 cells. The results obtained with the different constructs are presented in Figure 12. *I-SceI* induced DSBs clearly stimulate the SSA repair mechanism. Furthermore, homology of 70 bp was sufficient to achieve nearly maximum efficiencies of induced SSA, while the level of spontaneous recombination (without *I-SceI* induced DSB) was minimal. With duplication of 220 bp maximum efficiency was achieved while no additional gains in SSA efficiency were observed with longer duplications. Similar results were obtained with COS-7 cells (data not shown). 70 and 220 bp of homology gave the best ratio of activity vs background. Because β-galactosidase is assayed in cell lysates and one single cell can contain several copies of the target plasmid, it is impossible to evaluate the absolute SSA efficiency by this method. Therefore direct coloration of the cellular monolayer was performed 72 hours post-transfection (Figure 13). Virtually no blue cells were detected in the absence of the meganuclease (Figure 13A). In contrast, many β-galactosidase-positive cells are present when *I-SceI* is cotransfected with the target vector, demonstrating the stimulation of homologous recombination by meganuclease induced DSB. The efficiency of *I-SceI* induced SSA was calculated by counting the blue cells (cells where recombination has taken place) and comparing it with the number of transfected cells (cells that effectively received

DNA). Figure 13B shows that 50 to 60% of the cells undergo homologous recombination when *I-SceI* is present along with the target vector carrying 70 or 220 bp duplications while spontaneous recombination represents less than 0,1% of the events. Thus, the construct with the 70 bp and 220 bp of homology as well as the transgene were selected for the animal study.

B. Meganuclease-induced recombination of an extrachromosomal reporter *in toto*

- Experimental procedures

Hydrodynamic-based transfection in vivo

Transduction of the mouse liver cells was performed by hydrodynamic tail vein injections as previously described (Zhang, G., Budker, V., Wolff, A., 1999, Human Gene Therapy, 10, 1735-1737; Liu, F.,Song, Y.K., Liu, D., 1999, Gene Therapy, 6, 1258-1266). This method allows efficient transduction and expression of exogenous genes in animals by administration of plasmid DNA by tail vein injection. Briefly, DNA is mixed in 1.5 to 2 ml of PBS, which represents 10% of the animal's weight. Tail vein injections are subsequently performed with a 26-gauge needle over a 5-10 sec period using sterile materials and working conditions. Using such a protocol, almost exclusively liver cells are transduced, thus the I-*Sce*I-mediated SSA event leading to the correction of the *LagoZ* gene was studied in the liver. The I-*Sce*I expressing vector used is the pCLS 197 corresponding to the I-*Sce*I-coding sequences (Patent US 5,474,896) under the control of the CMV promoter in a pUC backbone and is 5737bp long.

OF1 mice weighing fifteen to twenty grams were obtained from Charles River Laboratories, France. A total of twenty micrograms of DNA, containing equal amounts of target vector and either an I-*Sce*I expression or control vector, was injected into mouse tail veins. The target vector contains the *LagoZ* gene interrupted by an I-*Sce*I cleavage site flanked by direct repeat sequences containing 70 bp of homology. Control mice were injected with a mixture of the target vector and a plasmid that does not express I-*Sce*I.

β-galactosidase activity

Three days after injection, mice were euthanized by cervical dislocation and X-Gal stainings of their livers were performed. Livers were dissected out of

the animals in cold 1x PBS and the lobes were cut in pieces of about one fourth a centimeter in order to allow a better access of the X-Gal in the tissue. Then liver pieces were placed in fresh cold PBS 1x in a 12-well cell culture plate kept on ice, and fixed in 4% paraformaldehyde for 1 hour under agitation at 4°C. Samples were then washed 3 times at room temperature for 30 minutes with wash buffer (100mM sodium phosphate pH=7.3, 2mM $MgCl_2$, 0.01% sodium deoxycholate, 0.02% NP-40 by volume). *In toto* X-Gal staining was performed overnight at 37°C in staining solution (5mM potassium ferricyanide, 5mM potassium ferrocyanide, 1mg/ml X-gal, 20mM Tris pH=7.3 in wash buffer). Finally samples were washed extensively with PBS and examined under microscope. Pictures were taken with a Nikon Coolpix camera under a Nikon SMZ 1500 binocular.

- Results

Cellular study has shown that homology of 70 bp is sufficient to achieve nearly maximal efficiencies of DSB induced SSA, while the level of spontaneous recombination (without I-*Sce*I induced DSB) is minimal. Thus, in a first attempt to stimulate recombination *in vivo*, transient experiments were performed. A mixture of the target vector (30 µg) and either the I-*Sce*I expression or control plasmid (10 µg) were introduced into the liver via a hydrodynamic tail vein injection method. Figure 12 shows a magnified picture of liver collected and stained 3 days after injection. Blue dots represent cells where a defective *LagoZ* gene, bearing an I-*Sce*I site flanked by a 70 bp duplication, was repaired. After meganuclease induced DSB, the SSA pathway results in the deletion of one repeat and reconstitution of a functional gene. Active ß-galactosidase encoded by the *LagoZ* gene can then be detected by X-Gal stainings. Furthermore, no gene correction was detected in the absence of the meganuclease expression vector. These data represent the first evidence that meganuclease induced recombination can be stimulated in liver and that *in toto* repair of an extrachromosomal target can be achieved.

Exemple 8: Meganuclease-induced recombination of a chromosomal reporter *in toto* using I-S*ce* I expressing adenovirus.

In order to demonstrate meganuclease-induced genomic surgery of a chromosomal reporter in toto in different mice tissues, the repair of the *lagoZ* gene *in toto* was tested by transducing cells of several organs with an I-S*ce*I-expressing

adenovirus, « Ad.I-*Sce*I ». Control transgenic littermates were infected with a non-I-SceI-expressing adenovirus, « Ad.control ». Adenovirus infections in transgenic mice were performed by intraveinous (IV) injections. Repair of the *lagoZ* gene *in toto* in several tissues was then tested by two methods that detect ß-galactosidase activity *in toto*, X-gal staining and FDG assays.

- Experimental procedures

Transgenic mice

The transgene used for the generation of transgenic founders was a *Bgl*II/*Not*I fragment of 5919bp carrying the defective *LagoZ* gene, inactivated by a *LagoZ* duplication of 70 bp or 220bp and the *I-SceI* cleavage site, under the control of the human elongation factor 1 alpha promoter (See Figure 11).

Transgenic founder were generated by classical transgenesis, i.e. by microinjecting the linear and purified *Bgl*II/*Not*I fragment described above at 500copies/picolitres into the male pronuclei of fertilized ova at the one-cell stage derived from the mating of B6D2F1 males and females purchased from Elevage Janvier. Microinjections were performed under a Nikon TE2000 microscope with Normarski DIC with eppendorf transferMan NK2 micromanipulators and eppendorf Femtojet 5247 micro-injector. After injections, ova were transferred to surrogate pseudopregnant B6CBAF1/J females (Elevage Janvier) for development and delivery. Transgenic mice generated by this procedure were identified by PCR and Southern Blot analysis on genomic DNA extracted from tail biopsies of F0 mice. The molecular characterization of the transgene integration was done by PCR and Southern Blot analysis.

Then the founder were mated to B6D2F1 mice in order to obtain hemizygote transgenic F1 animal. Expression of the transgene was tested by performing an RT-PCR experiment on RNAs extracted from a tail biopsie from a transgenic F1 animal using Qiagen RNeasy kit (cat N°74124). Hemizygote F1 mice were then mated to B6D2F1 mice in order to establish an F2 hemizygote transgenic strain.

Two independent strains were used bearing either 220bp or 70bp long *lagoZ* gene repeated sequences. These transgenic strains are referred as strain « 361 » and « 58A», respectively.

The molecular characterization of the transgene integration showed that the integration is about 5 direct repeats of the *Bgl*II/*Not*I transgene in « 361 » and 2 inverted repeats plus 5 direct repeats in « 58A ». Hemizygote mice were identified by tail biopsies, genomic DNA extraction and PCR analysis. « 361 » and « 58A» hemizygote mice were then used for *in toto* I-*Sce*I mediated-*lagoZ* gene repair and transgenic littermates were used as negative controls.

Adenovirus-based transduction in toto

Recombinant type V adenovirus bearing the I-*Sce*I meganuclease coding region under the control of a CMV promoter, « Ad.I-*Sce*I » was provided by Q BIO gene company at 1.58 10^{11} infectious units concentration scored by the $TCID_{50}$ method. The negative adenovirus control « Ad.control » was as well provided by Q BIO gene company at 3.76 10^{11} infectious units concentration. Recombinant type V adenovirus infections were performed by intraveinous (IV) injections in transgenic mice tail veins. Transgenic mice were weighed and anesthetized before infections by intraperitoneal injection of a mixture of Xylasin (100 mg/kg) and Ketamine (10 mg/kg). IV infections were performed with 10^{10} infectious units/animal in a volume of 400 µl. Infections were performed in 4 to 7 weeks-old transgenic mice. Adenovirus-infected mice and uninfected control littermates were bred in isolator untill sacrificed for β-galactosidase assays.

β-galactosidase activity

Adenovirus-infected mice were sacrificed by $C0_2$ inhalation from 5 to 14 days-post-infections (dpi) and their organs were processed for β-galactosidase assays. About 10% of the liver (8 mm^3) was emplyed for protein extraction and the remaining 90% was used for β-galactosidase in toto X-gal assays (protocol described previously).

Fluorescent β-galactosidase assays were incubated at 37°C in 96 well plate. The assays were performed in a total volume of 100 µl containing 30 µl of protein extract, 1 µM Fluorescein digalactoside (FDG, Sigma), 0.1% β-mercaptoethanol, 1 mM $MgCl_2$, and 100 mM Phosphate buffer, pH 7.0. The plates were scanned on the Fuoroskan Ascent (Labsystem) at 5-minutes intervals. The β-galactosidase activity is calculated as relative unit normalized for protein concentration and incubation time.

- Results

Two « 58A » transgenic mice were IV-injected with 10^{10} infectious units of « Ad.I-*Sce*I» adenovirus in order to target a DSB in-between the 70bp duplicated *lagoZ* sequences and induce the repair of the reporter gene. At various times post-injection the mice were sacrificed and several organs were dissected and analyzed by *in toto* X-gal assays. Blue staining was detected as dispersed cells over the entire liver of infected mouse euthanized at 5dpi (Figure 15A). No staining could be detected in the other organs tested, i.e. kidneys, spleen, heart and lungs. Two « 58A » transgenic mouse littermates were used as controls, one IV-injected with 10^{10} infectious units of the control adenovirus « Ad.control » and the other uninfected. No β -galactosidase activity could be detected in the liver of either control (data not shown). Similar results were obtained with two « 361 » transgenic mice injected with 10^{9} and 10^{10} infectious units of the Ad.I-*Sce*I adenovirus (Figure 15B and 15C respectively). These results were confirmed by measuring the β-galactosidase activity in liver extract (Figure 16). A high activity was detected in liver of mice injected with Adenovirus expressing I-SceI (Ad.1Sce-I). In contrast, Non-injected mice (NI) shows only a residual background activity similar to the activity detected in mice injected with the control adenovirus (Ad.control).

The IV-injected mouse with 10^{10} infectious units of «Ad.1-SceI» adenovirus exhibited more stained liver cells and more β-galactosidase activity than the IV-injected mouse with 10^{9} infectious units of «Ad.1-SceI» adenovirus. These results suggest that I-*Sce*I-induced recombination could be dose dependent and that a better yield of I-SceI induced recombination could be obtained by increasing the injected-adenovirus titer. Thus, I-SceI-induced genome surgery should be detectable in other organs reported to be less sensitive to type V adenovirus infection.

Taken together, these data strongly suggest that the reporter gene repair was induced by the activity of the I-*Sce*I meganuclease. This result is the first evidence that I-S*ce*I and more generally the meganucleases can be used *in toto* to induce efficient site-specific homologous recombination leading to the repair of a chromosomal gene. Thus, this result opens applications in the field of gene therapy in mammals.

CLAIMS

1°) A method for producing a meganuclease variant derived from an initial meganuclease, said meganuclease variant being able to cleave a DNA target sequence which is different from the recognition and cleavage site of the initial meganuclease, characterized in that it comprises the following steps:

a) preparing a library of meganuclease variants from an initial meganuclease ; and,

b) isolating the variants able to cleave said DNA target sequence.

2°) The method according to claim 1, characterized in that said meganuclease variants have amino acid variations at positions contacting the DNA target or interacting directly or indirectly with said DNA target.

3°) The method according to any one of claims 1 or 2, characterized in that said initial meganuclease is a natural or a modified meganuclease.

4°) The method according to any one of claims1 to 3, characterized in that said initial meganuclease is a homing endonuclease.

5°) The method according to claim 4, characterized in that said homing endonuclease is a LAGLIDADG homing endonuclease.

6°) The method according to claim 5, characterized in that said LAGLIDADG homing endonuclease is selected from the group consisting of: I-Cre I, I-Dmo I, PI-Sce I, and PI-Pfu I.

7°) The method according to claim 6, characterized in that said LAGLIDADG homing endonuclease is I-Cre I.

8°) The method according to claim 2, characterized in that said amino acid variations are replacement of the initial amino acid with an amino acid selected from the group consisting of: D, E, H, K, N, Q, R, S, T, Y.

9°) The method according to claim 7, characterized in that said library of I-Cre I variants is prepared by introducing amino acid diversity in positions

selected from the group consisting of: Q26, K28, N30, S32, Y33, Q38, Q44, R68, R70 and T140.

10°) The method according to claim 9, characterized in that said library of I-Cre I variants is prepared by introducing amino acid diversity in positions : a) Q26, K28, N30, Y33, Q38, Q44, R68, R70, T140; b) Q26, K28, N30, Y33, Q38, Q44, R68, R70; c) Q26, K28, N30, Y33, Q44, R68, R70 ; or d) Q26, K28, Y33, Q38, Q44, R68, R70.

11°) The method according to claim 10, characterized in that said library of I-Cre I variants is prepared by introducing amino acid diversity in positions: Q26, K28, N30, Y33, Q38, Q44, R68 and R70.

12°) The method according to anyone of claims 1 to 11, characterized in that said isolation of the variants able to cleave the DNA target sequence comprises the steps selected in the group consisting of:

a) a selection step for the binding ability, a screening step for the binding ability, a selection step for the cleavage activity, and a screening step for the cleavage activity;

b) a selection step for the binding ability, a screening step for the binding ability, and a screening step for the cleavage activity;

c) a selection step for the binding ability, a selection step for the cleavage activity, and a screening step for the cleavage activity; or,

d) a screening step for the binding ability and a screening step for the cleavage activity.

13°) The method according 12, characterized in that it comprises or consists of the following steps: a selection step for the binding ability, a selection step for the cleavage activity, and a screening step for the cleavage activity.

14°) The method according 12, characterized in that said selection and screening steps for the binding ability use the phage display.

15°) The method according 12, characterized in that said selection for the cleavage activity uses a test in which the cleavage leads to either the activation of a positive selection marker or the inactivation of a negative selection marker.

16°) The method according to daim 12, characterized in that said screening for the cleavage activity uses a test in which the cleavage leads to: a) the activation of a positive selection marker or a reporter gene; or b) the inactivation of a negative selection marker or a reporter gene.

17°) The method according to claim 15 or 16, characterized in that : said selection for the cleavage activity uses a test in which the cleavage leads to the activation of a positive selection marker, and said screening for the cleavage activity uses a test in which the cleavage leads to the activation of a a reporter gene.

18°) A meganuclease variant, characterized in that it is presenting a recognition and cleavage site different from the site of the initial meganuclease from which it is derived.

19°) The meganuclease according to claim 18, characterized in that it is obtainable by a method according to anyone of claims 1 to 17.

20°) A polynucleotide, characterized in that it is encoding a custom-made meganuclease according to claim 18 or 19.

21°) A vector, characterized in that it comprises a polynucleotide according to claim 18 or 19.

22°) The vector according to claim 21, characterized in that it further comprises a targeting DNA construct.

23°) The vector according to claim 21 or 22, characterized in that said targeting DNA construct comprises sequences sharing homologies to the region surrounding the cleavage site of the meganuclease according to claim 18 or 19.

24°) The vector according to claim 21 or 22, characterized in that said targeting DNA construct comprises:

a) sequences sharing homologies to the region surrounding the cleavage site of the meganuclease according to claim 18 or 19, and

b) sequences to be introduced flanked by sequences as in a).

25°) A cell, characterized in that it is modified by a polynucleotide according to claim 20 or a vector according to anyone of claims 21 to 24.

26°) A transgenic plant, characterized in that it comprises a polynucleotide according to claim 20 or a vector according to anyone of claims 21 to 24.

27°) A non-human transgenic mammal, characterized in that it comprises a polynucleotide according to claim 20 or a vector according to anyone of claims 21 to 24.

28°) Use of a meganuclease according to claim 18 or 19, a polynucleotide according to claim 20, a vector according to anyone of claims 21 to 24, a cell according to claim 25, a transgenic plant according to claim 26, a non-human transgenic mammal according to claim 27, for molecular biology, for *in vivo* or *in vitro* genetic engineering, and for *in vivo* or *in vitro* genome engineering.

29°) The use according to claim 28, for inducing a double-stranded break in a site of interest comprising a DNA target sequence, thereby inducing a DNA recombination event, a DNA loss or cell death.

30°) The use according to claim 29, characterized in that said double-stranded break it is for: repairing a specific sequence, modifying a specific sequence, restoring a functional gene in place of a mutated one, attenuating or activating an endogenous gene of interest, introducing a mutation into a site of interest, introducing an exogenous gene or a part thereof, inactivating or deteting an endogenous gene or a part thereof, translocating a chromosomal arm, or leaving the DNA unrepaired and degraded.

31°) The use according to anyone of claims 28 to 30, characterized in that said meganuclease, polynucleotide, vector, cell, transgenic plant or non-human transgenic mammal are associated with a targeting DNA construct as defined in claim 23 or 24.

32°) A method of genetic engineering, characterized in that it comprises the step of introducing, with a meganuclease according to claim 18 or 19, a

double-stranded break in a site of interest located on a vector and comprising said DNA target sequence, thereby inducing a homologous recombination with another vector presenting homology with the sequence surrounding the cleavage site.

33°) A method of genome engineering, characterized in that it comprises the following steps: 1) introducing a double-stranded break at the genomic locus comprising at least one recognition and cleavage site of a meganuclease according to claim 18 or 19; 2) providing a targeting DNA construct comprising the sequence to be introduced flanked by sequences sharing homologies to the targeted locus.

34°) A method of genome engineering, characterized in that it comprises the following steps: 1) introducing a double-stranded break at the genomic locus comprising at least one recognition and cleavage site of a a meganuclease according to claim 18 or 19; 2) maintaining under conditions appropriate for homologous recombination with chromosomal DNA sharing homologies to regions surrounding the cleavage site.

35°) A composition, characterized in that it comprises at least one meganuclease, polynucleotide or vector according to anyone of claims 18 to 24.

36°) The composition according to claim 35, characterized in that said composition further comprises a targeting DNA construct comprising the sequence which repairs the site of interest flanked by sequences sharing homologies to the targeted locus.

37°) Use of a at least one meganuclease, polynucleotide or vector according to anyone of claims 18 to 24, for the preparation of a medicament for preventing, improving or curing a genetic disease in an individual in need thereof, said medicament being administered by any means to said individual.

38°) Use of a at least one meganuclease, polynucleotide or vector according to anyone of claims 18 to 24, for the preparation of a medicament for preventing, improving or curing a disease caused by an infectious agent that presents a DNA intermediate, in an individual in need thereof, said medicament being administered by any means to said individual.

39°) Use of a at least one meganuclease, polynucleotide or vector according to anyone of claims 18 to 24, *in vitro*, for inhibiting the propagation, inactivating or deleting an infectious agent that presents a DNA intermediate, in biologically derived products or products intended for biological uses or for disinfecting an object.

40°) The use according to claim 39, characterized in that said infectious agent is a virus.

Amino acid sequence of a Single chain I-Cre I (SEQ ID NO: 1)

```
  1 MANTKYNKEF LLYLAGFVDG DGSIIAQIKP NQSYKFKHQL SLTFQVTQKT
 51 QRRWFLDKLV DEIGVGYVRD RGSVSDYILS EIKPLHNFLT QLQAMLERIR
101 LFNMREFLLY LAGFVDGDGS IIAQIKPNQS YKFKHQLSLT FQVTQKTQRR
151 WFLDKLVDEI GVGYVRDRGS VSDYILSEIK PLHNFLTQLQ PFLKLKQKQA
201 NLVLKIIEQL PSAKESPDKF LEVCTWVDQI AALNDSKTRK TTSETVRAVL
251 DSLSEKKKSS PAAD
```

Polynucleotide sequence encoding the Single chain I-Cre I (SEQ ID NO: 2)

```
  1 ATGGCCAACA CTAAGTACAA TAAAGAATTT CTCCTGTATC TGGCAGGTTT
 51 CGTCGACGGC GATGGCTCCA TTATCGCACA GATCAAGCCG AATCAGAGCT
101 ACAAGTTTAA ACACCAACTG TCTCTCACTT TCCAGGTTAC CCAGAAAACT
151 CAACGTCGCT GGTTCCTGGA TAAGCTGGTA GATGAGATCG GTGTGGGCTA
201 TGTACGCGAC CGTGGCTCTG TGAGCGACTA TATCCTGTCT GAGATTAAAC
251 CACTGCATAA TTTTCTGACC CAGCTGCAGG CTATGCTGGA GCGTATCCGT
301 CTGTTCAACA TGCGTGAGTT CCTGCTGTAC CTGGCCGGCT TTGTGGACGG
351 TGACGGTAGC ATCATCGCTC AGATTAAACC AAACCAGTCT TATAAATTCA
401 AGCATCAGCT GTCCCTGACC TTTCAGGTGA CTCAAAAGAC CCAGCGCCGT
451 TGGTTTCTGG ACAAACTGGT GGATGAAATT GGCGTTGGTT ACGTACGTGA
501 TCGCGGTAGC GTTTCCGATT ACATTCTGAG CGAAATCAAG CCGCTGCACA
551 ACTTCCTGAC TCAACTGCAA CCGTTTCTGA AACTGAAACA GAAACAGGCA
601 AACCTGGTTC TGAAAATTAT CGAACAGCTG CCGTCTGCAA AAGAATCCCC
651 GGACAAATTC CTGGAAGTTT GTACCTGGGT GGATCAGATT GCAGCTCTGA
701 ACGATTCTAA GACGCGTAAA ACCACTTCTG AAACCGTTCG TGCTGTGCTG
751 GACAGCCTGA GCGAGAAGAA GAAATCCTCC CCGGCGGCCG ACTAG
```

Figure 1

FIGURE 2A (SEQ ID NO: 3)

```
1   ATGGCCAATA CCAAATATAA CAAAGAGTTC CTGCTGTACC TGGCCGGCTT
51  TGTGGACGGT GACGGTAGCA TCATCGCTCA GATTAAACCA AACCAGTCTT
101 ATAAATTCAA GCATCAGCTG TCCCTGACCT TTCAGGTGAC TCAAAAGACC
151 CAGCGCCGTT GGTTTCTGGA CAAACTGGTG GATGAAATTG GCGTTGGTTA
201 CGTACGTGAT CGCGGTAGCG TTTCCGATTA CATTCTGAGC GAAATCAAGC
251 CGCTGCACAA CTTCCTGACT CAACTGCAAC CGTTTCTGAA ACTGAAACAG
301 AAACAGGCAA ACCTGGTTCT GAAAATTATC GAACAGCTGC CGTCTGCAAA
351 AGAATCCCCG GACAAATTCC TGGAAGTTTG TACCTGGGTG GATCAGATTG
401 CAGCTCTGAA CGATTCTAAG ACGCGTAAAA CCACTTCTGA AACCGTTCGT
451 GCTGTGCTGG ACAGCCTGAG CGAGAAGAAG AAATCCTCCC CG
```

FIGURE 2B (SEQ ID NO: 4)

```
1   ATGGCCAACA CTAAGTACAA TAAAGAATTT CTCCTGTATC TGGCAGGTTT
51  CGTCGACGGC GATGGCTCCA TTATCGCACA GATCAAGCCG AATCAGAGCT
101 ACAAGTTTAA ACACCAACTG TCTCTCACTT TCCAGGTTAC CCAGAAAACT
151 CAACGTCGCT GGTTCCTGGA TAAGCTGGTA GATGAGATCG GTGTGGGCTA
201 TGTACGCGAC CGTGGCTCTG TGAGCGACTA TATCCTGTCT GAGATTAAAC
251 CACTGCATAA TTTTCTGACC CAGCTGCAGC CGTTCCTCAA GCTGAAGCAA
301 AAACAGGCCA ATCTCGTGCT GAAGATCATT GAGCAACTGC CATCCGCCAA
351 AGAGTCTCCG GATAAATTTC TGGAGGTCTG CACTTGGGTT GACCAAATCG
401 CTGCACTCAA CGACTCCAAA ACCCGCAAGA CGACCAGCGA GACTGTACGC
451 GCAGTTCTGG ATTCTCTCTC CGAAAAAAAG AAGTCTAGCC CG
```

FIGURE 2C (SEQ ID NO: 5)

```
1   ATGGCCAATA CCAAATATAA CAAAGAGTTC CTGCTGTACC TGGCCGGCTT
51  TGTGGACGGT GACGGTAGCA TCATCGCTCA GATTAAACCA AACCAGTCTT
101 ATAAGTTTAA ACATCAGCTA AGCTTGACCT TTCAGGTGAC TCAAAAGACC
151 CAGCGCCGTT GGTTTCTGGA CAAACTAGTG GATGAAATTG GCGTTGGTTA
201 CGTACGTGAT CGCGGATCCG TTTCCAACTA CATCTTAAGC GAAATCAAGC
251 CGCTGCACAA CTTCCTGACT CAACTGCAGC CGTTTCTGAA ACTGAAACAG
301 AAACAGGCAA ACCTGGTTCT GAAAATTATC GAACAGCTGC CGTCTGCAAA
351 AGAATCCCCG GACAAATTCC TGGAAGTTTG TACCTGGGTG GATCAGATTG
401 CAGCTCTGAA CGATTCTAAG ACGCGTAAAA CCACTTCTGA AACCGTTCGT
451 GCTGTGCTGG ACAGCCTGAG CGAGAAGAAG AAATCCTCCC CG
```

FIGURE 2D

UlibIfor (SEQ ID NO: 6)

ACGACGGCCA GTGAATTCAC CATGGCCAAT ACCAAATATA AC

UlibIrev (SEQ ID NO: 7)

1 CACCTGAAAG GTCAAGCTTA GMBBATGTTT AAACTTMBBA GACTGMBBTG

51 GMBBAATMBB AGCGATGATG CTACC

UlibIIfor (SEQ ID NO: 8)

GTTTAAACAT CAGCTAAGCT TGACCTTTVV KGTGACTCAA AAGACCCAG

UlibIIrev (SEQ ID NO: 9)

GATGTAGTTG GAAACGGATC CMBBATCMBB TACGTAACCA ACGCC

V = A or G or C

M = A or C

B = G or C or T

K = G or T

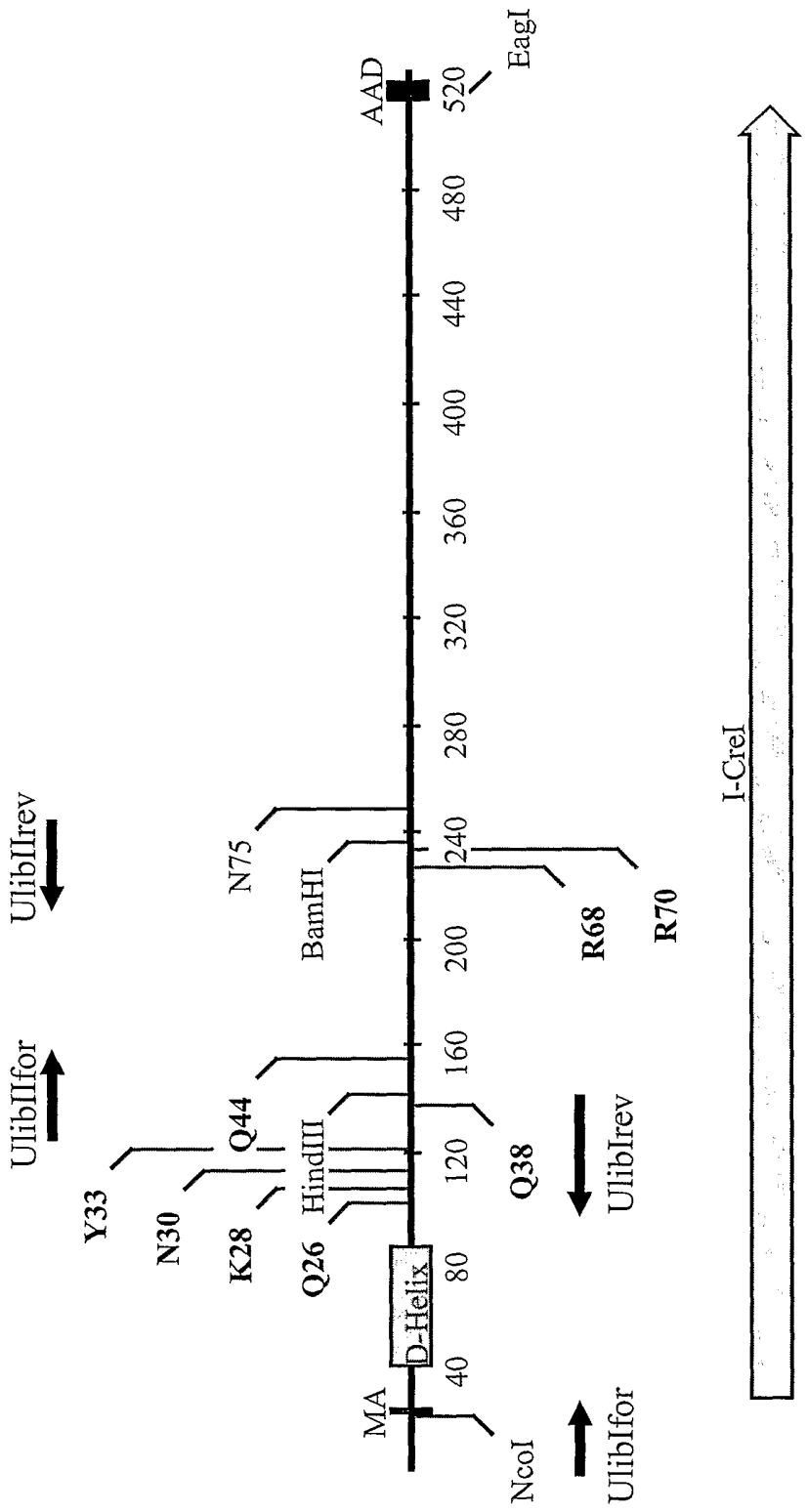

UlibIIrev
UlibIIfor
Y33
N30
K28
Q26
Q44
HindIII
N75
BamHI
Q38
R68
R70
UlibIrev
UlibIfor
D-Helix
MA
NcoI
AAD
EagI
40
80
120
160
200
240
280
320
360
400
440
480
520
I-CreI

FIGURE 3

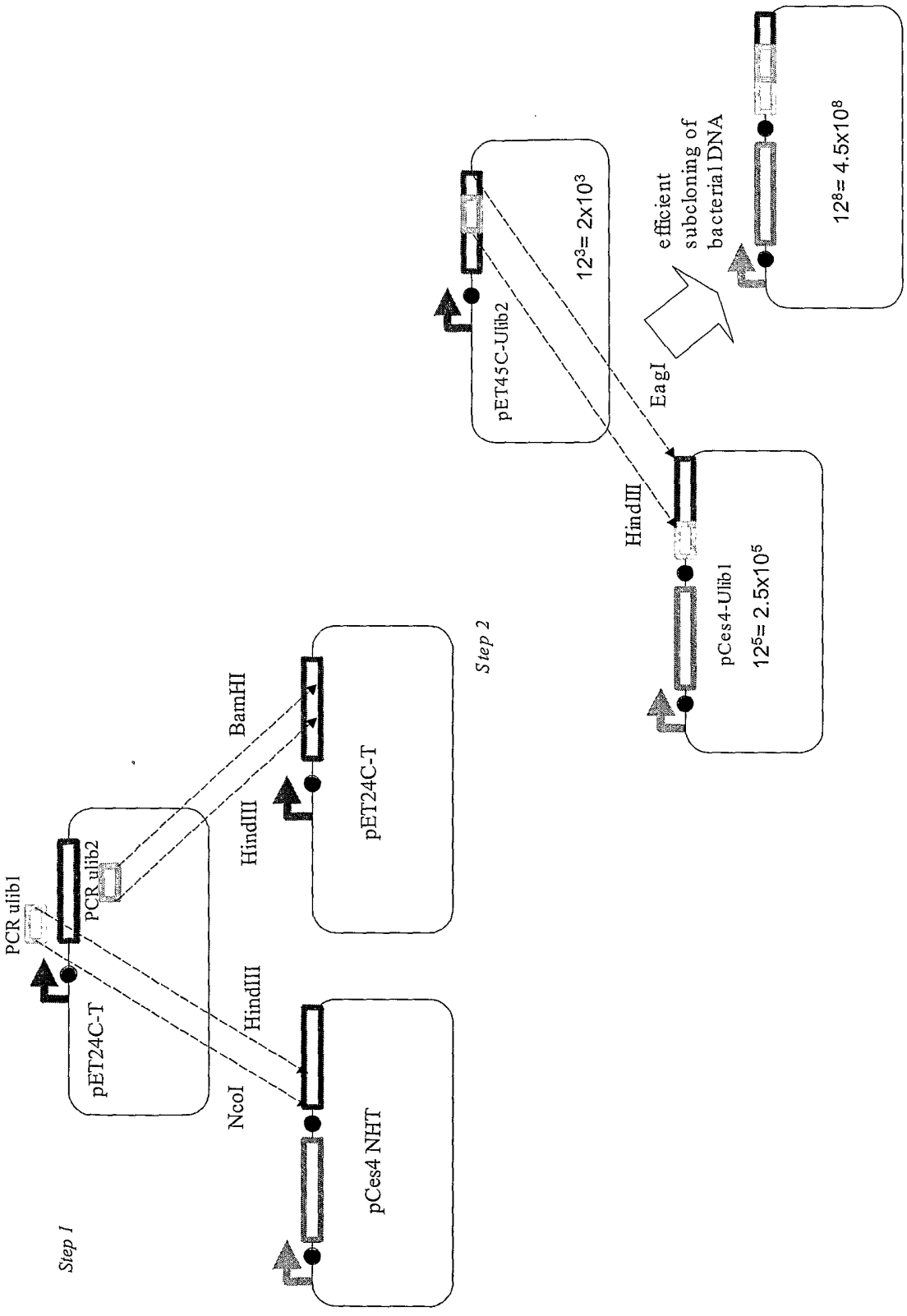
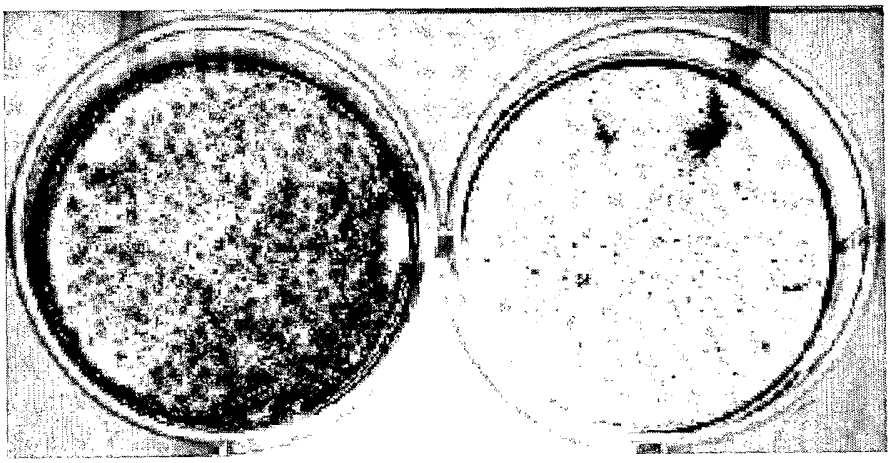

Step 1
pCes4 NHT
pET24C-T
PCR ulib1
PCR ulib2
NcoI
HindIII
HindIII
BamHI
pET24C-T
Step 2
pET45C-Ulib2
$12^3 = 2x10^3$
EagI
HindIII
pCes4-Ulib1
$12^5 = 2.5x10^5$
efficient subcloning of bacterial DNA
$12^8 = 4.5x10^8$

FIGURE 4



negative control

FIGURE 5A

I-Sce-I

FIGURE 5B

FIGURE 6A

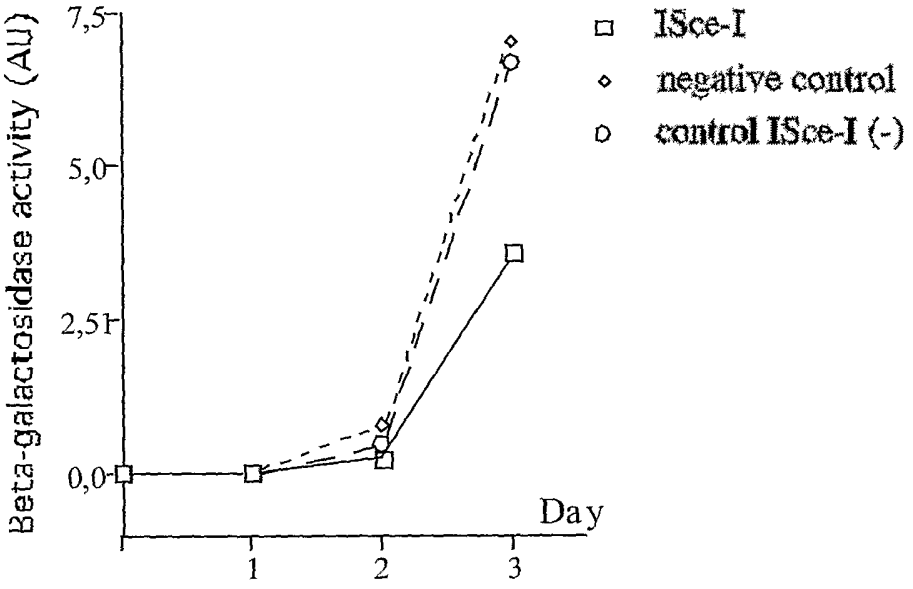


FIGURE 6B

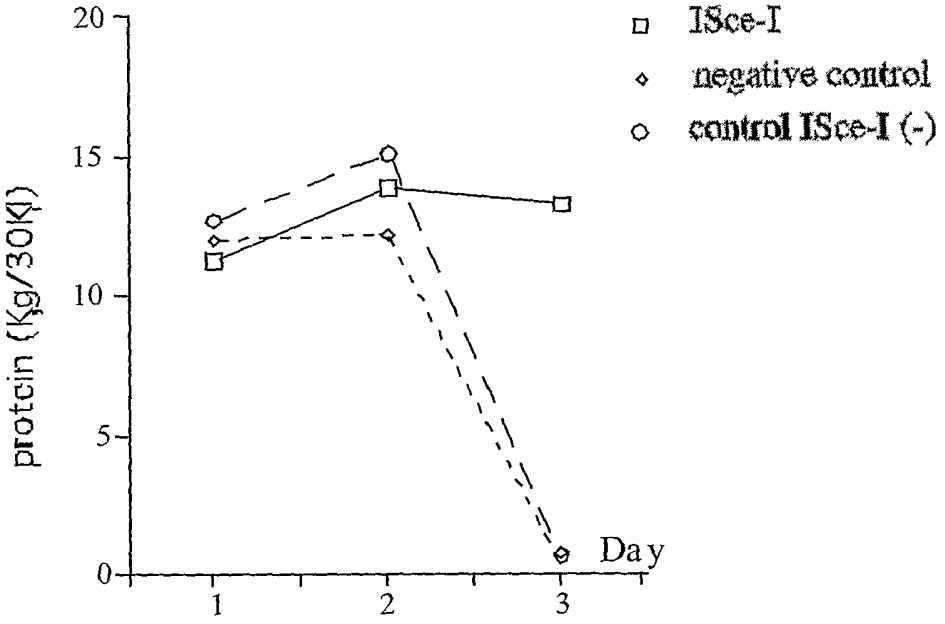


FIGURE 7A

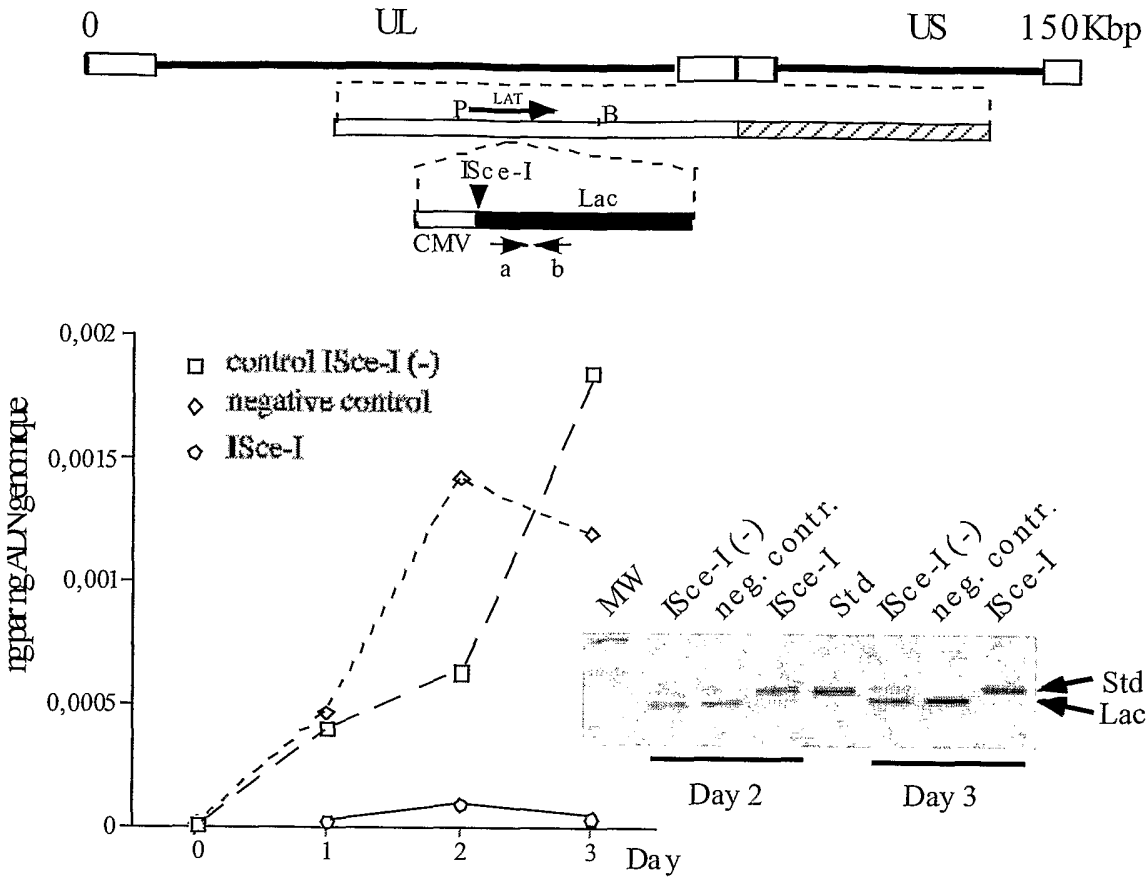
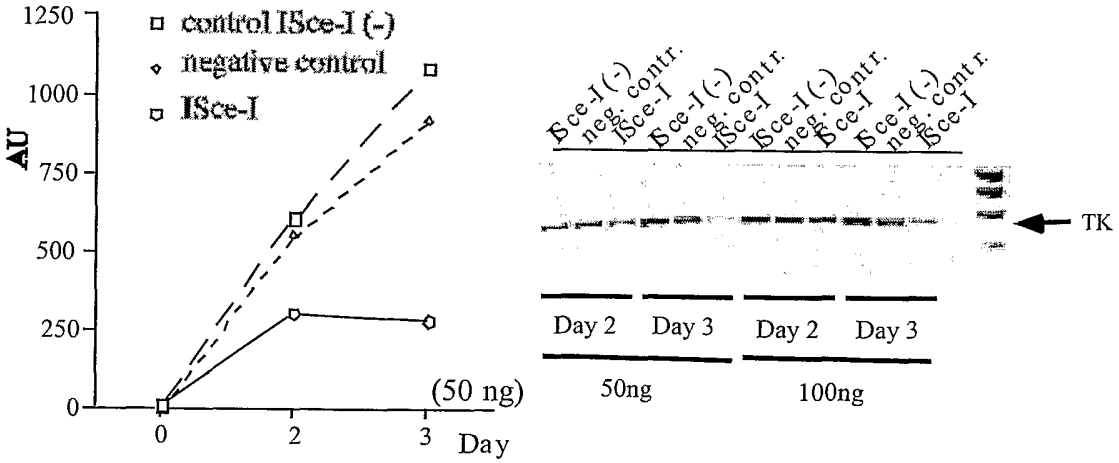


FIGURE 7B



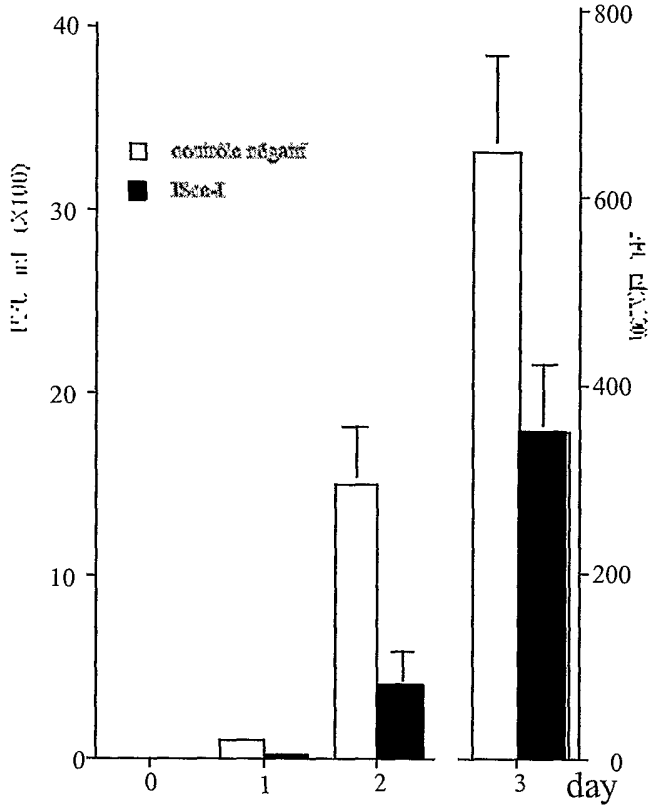


FIGURE 8

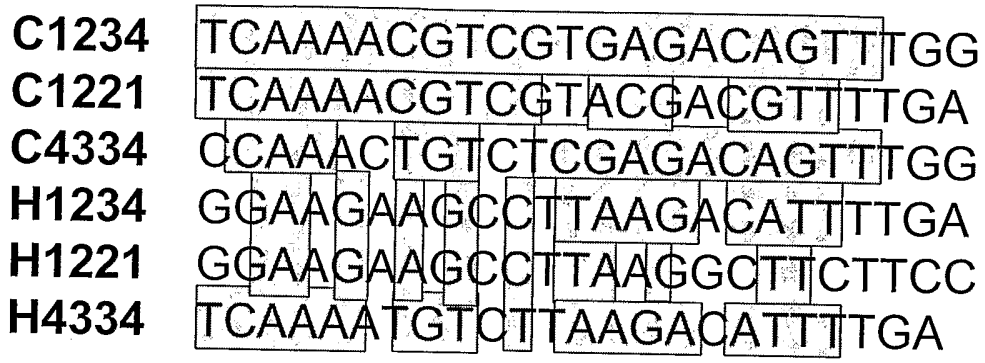

C1234 TCAAAACGTCGTGAGACAGTTTGG
C1221 TCAAAACGTCGTACGACGTTTTGA
C4334 CCAAACTGTCTCGAGACAGTTTGG
H1234 GGAAGAAGCCTTAAGACATTTTGA
H1221 GGAAGAAGCCTTAAGGCTTCTTCC
H4334 TCAAAATGTCTTAAGACATTTTGA

Figure 9

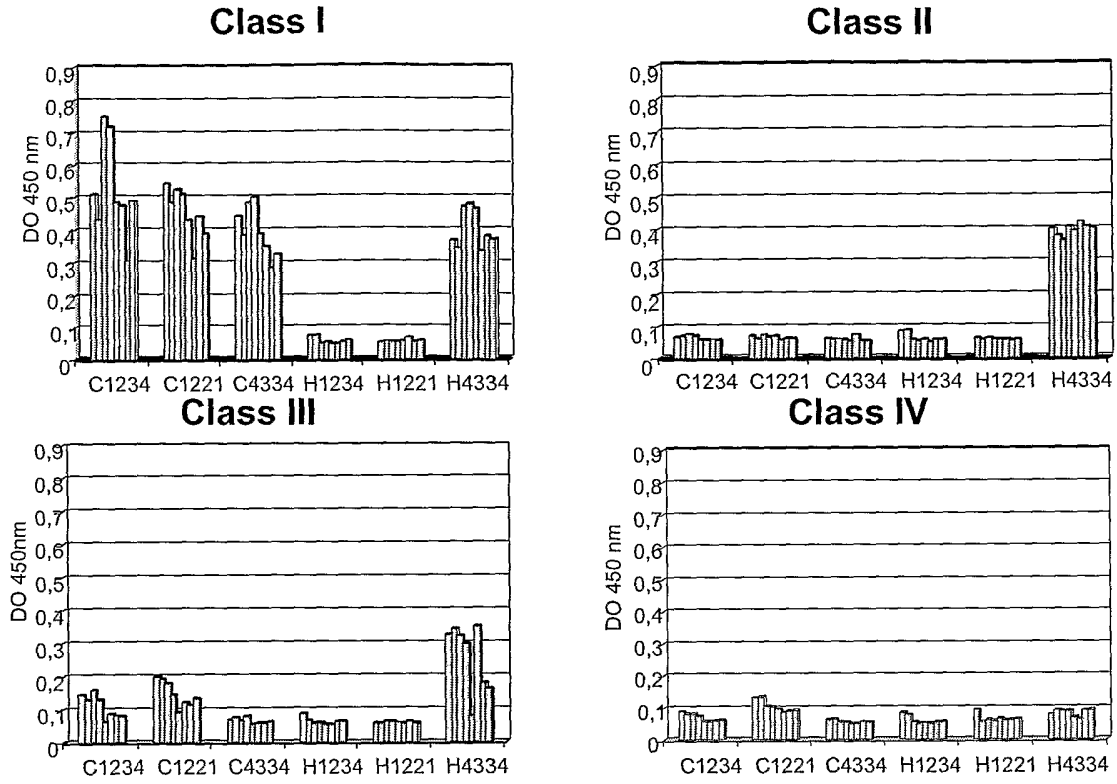

Class I
DO 450 nm
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
C1234 C1221 C4334 H1234 H1221 H4334
Class II
DO 450 nm
C1234 C1221 C4334 H1234 H1221 H4334
Class III
DO 450nm
C1234 C1221 C4334 H1234 H1221 H4334
Class IV
DO 450 nm
C1234 C1221 C4334 H1234 H1221 H4334

Figure 10

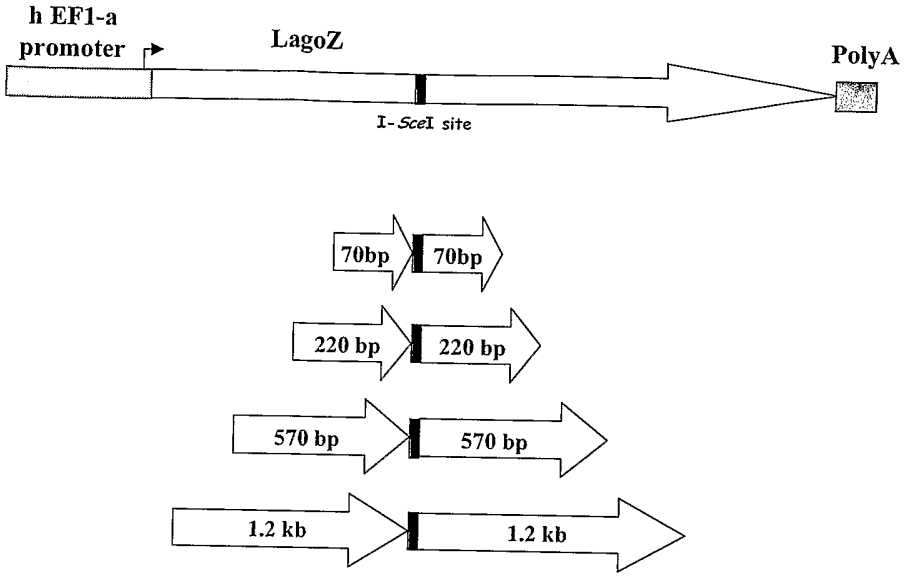

h EF1-a
promoter
LagoZ
PolyA
I-SceI site
70bp
70bp
220 bp
220 bp
570 bp
570 bp
1.2 kb
1.2 kb

Figure 11

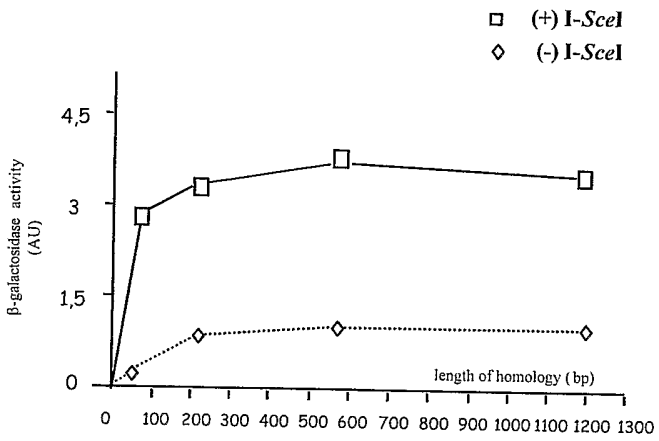
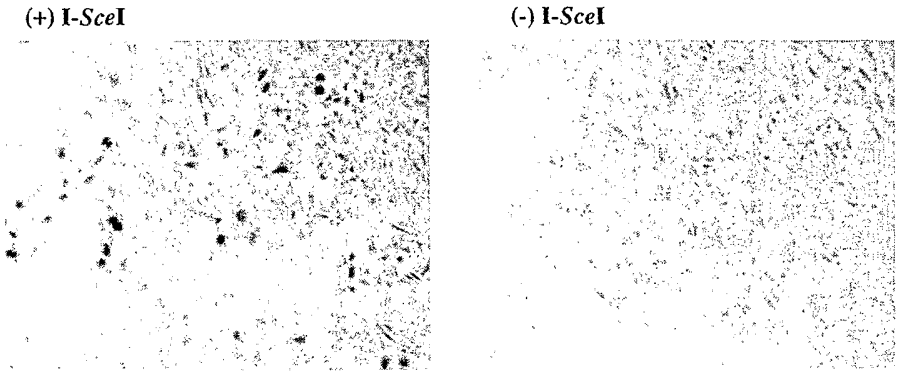

(+) I-SceI
(-) I-SceI
β-galactosidase activity (AU)
4,5
3
1,5
0
length of homology (bp)
0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300

Figure 12

A

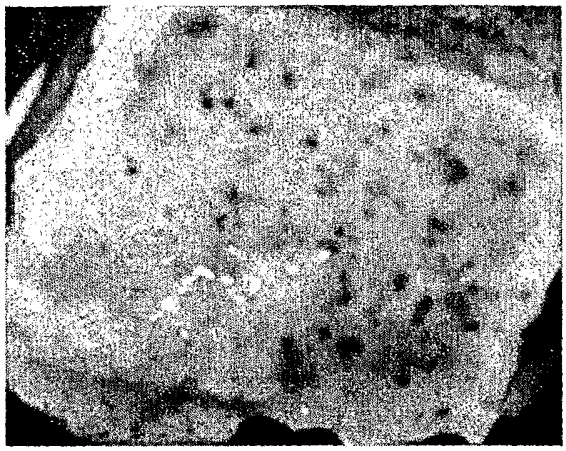


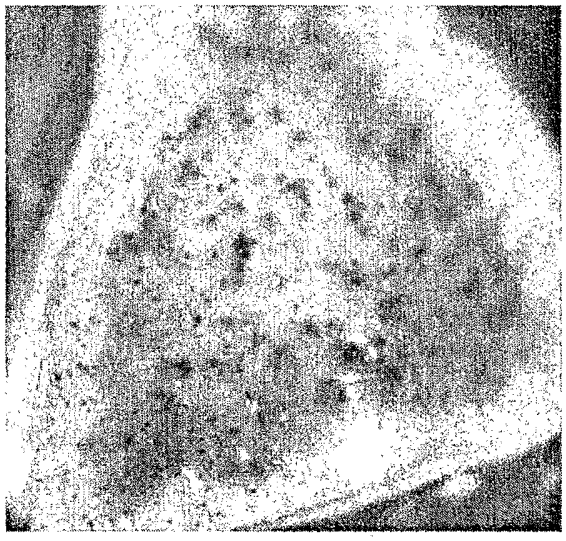
B

	Frequency of β - Gal $^+$ cells (per tranfected cells)	
Construct	(+) I-*Sce*I	(-) I-*Sce*I
SSA 70	60%	<0,1%
SSA 220	50%	<0,1%

FIGURE 13

Figure 14A





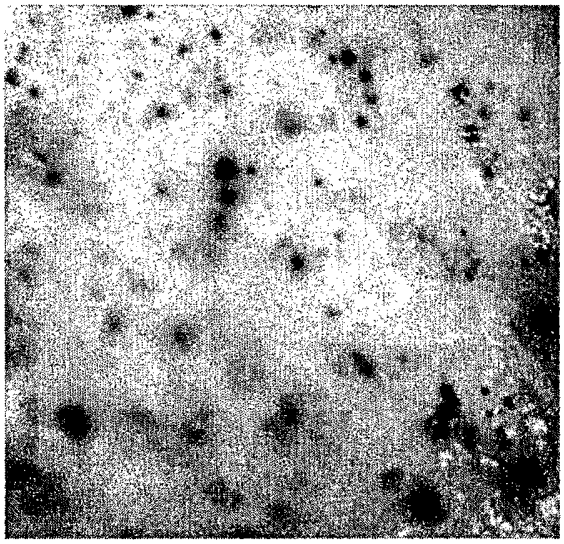
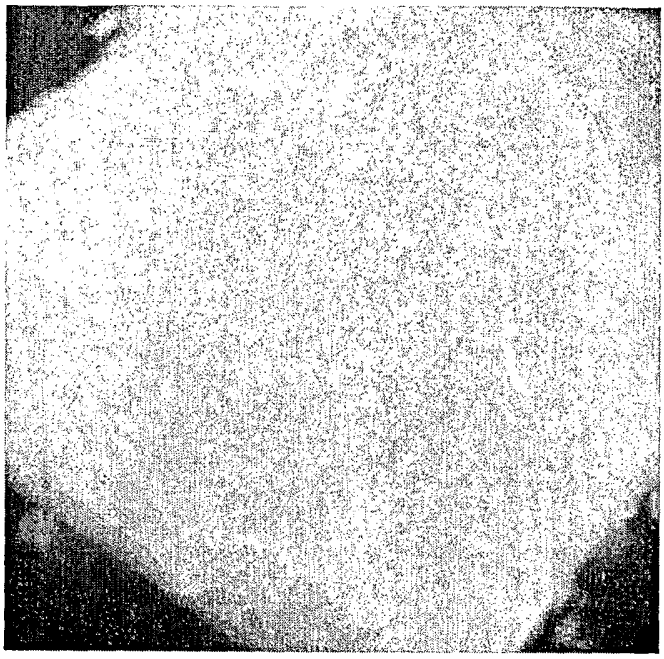


Figure 14B



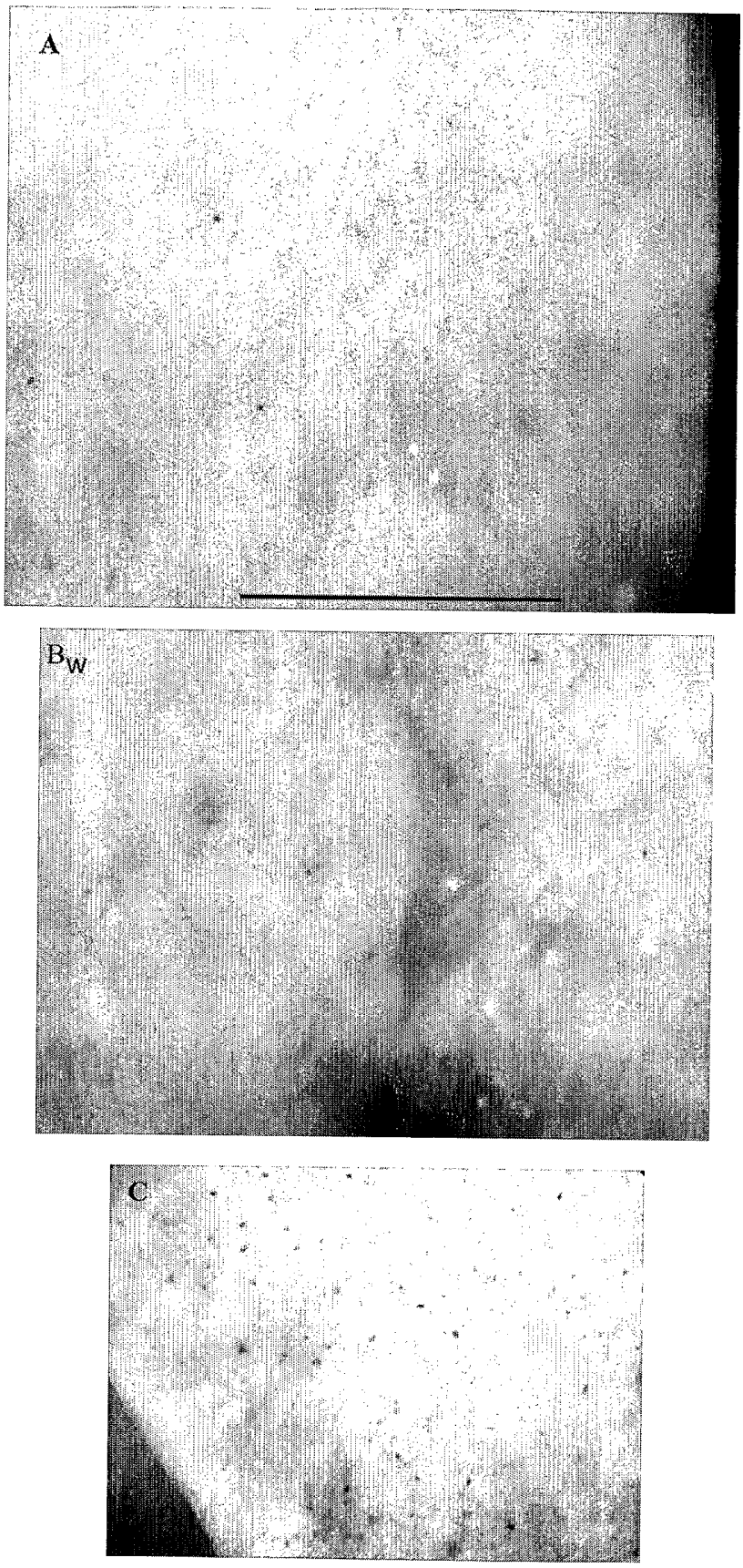


Figure 15

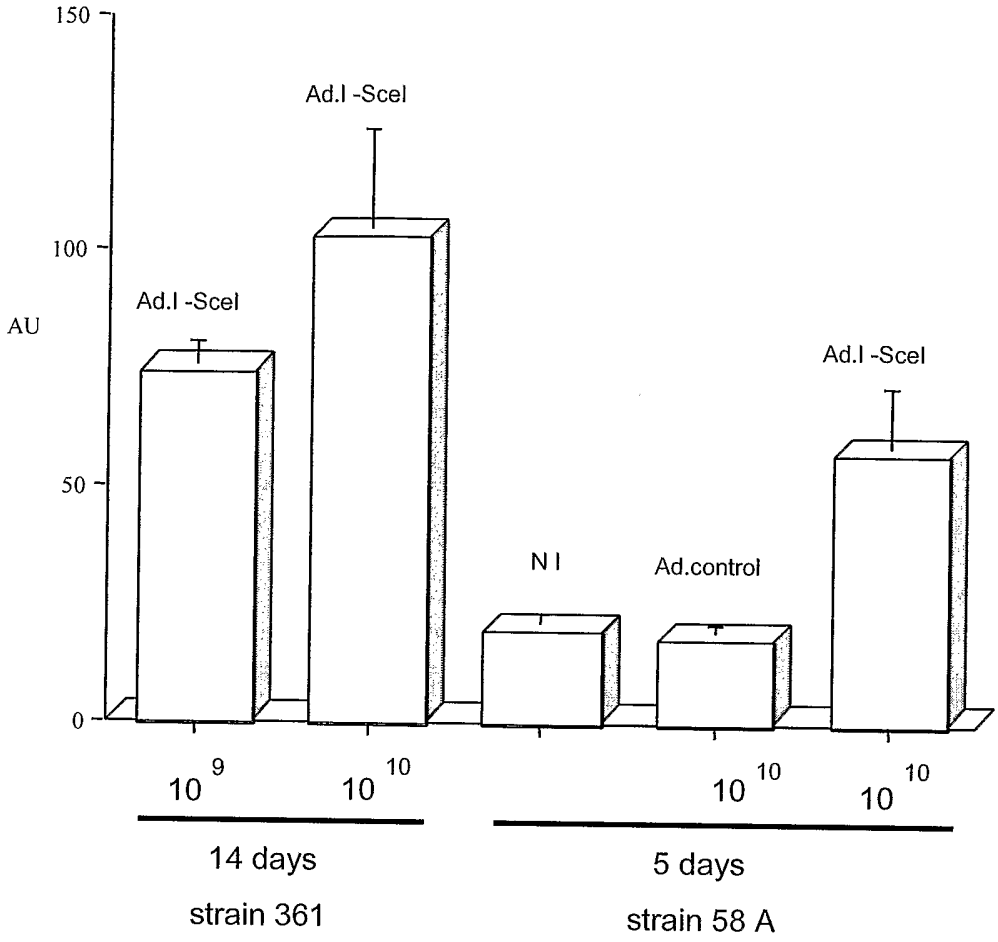

150
100
50
0
AU
Ad.I -SceI
Ad.I -SceI
Ad.I -SceI
N I
Ad.control
10 9
10 10
10 10
10 10
14 days
strain 361
5 days
strain 58 A

Figure 16

s1546PCT2.ST25
SEQUENCE LISTING

<110> CELLECTIS
ARNOUD, Sylvain
BRUNEAU, Sylvia
CABANIOLS, Jean-Pierre
CHAMES, Patrick
CHOULIKA, André
DUCHATEAU, Philippe
EPINAT, Jean-Pierre
GOUBLE, Agnès
LACROIX, Emmanuel
PAQUES, Frédéric
PEREZ-MICHAUT, Christophe
SMITH, Julianne
SOURDIVE, David

<120> Custom-made meganuclease and use thereof

<130> 1546PCT2

<150> US 60/491,535
<151> 2003-08-01

<150> US 60/442,911
<151> 2003-01-28

<160> 9

<170> PatentIn version 3.1

<210> 1
<211> 264
<212> PRT

s1546PCT2.ST25

<213> artificial sequence

<220>

<223> single chain I-Cre I

<400> 1

```
Met Ala Asn Thr Lys Tyr Asn Lys Glu Phe Leu Leu Tyr Leu Ala Gly
1               5                  10                  15

Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn Gln
            20                  25                  30

Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr Gln
        35                  40                  45

Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile Gly
    50                  55                  60

Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu Ser
65                  70                  75                  80

Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Ala Met Leu
                85                  90                  95

Glu Arg Ile Arg Leu Phe Asn Met Arg Glu Phe Leu Leu Tyr Leu Ala
            100                 105                 110

Gly Phe Val Asp Gly Asp Gly Ser Ile Ile Ala Gln Ile Lys Pro Asn
        115                 120                 125

Gln Ser Tyr Lys Phe Lys His Gln Leu Ser Leu Thr Phe Gln Val Thr
    130                 135                 140

Gln Lys Thr Gln Arg Arg Trp Phe Leu Asp Lys Leu Val Asp Glu Ile
145                 150                 155                 160

Gly Val Gly Tyr Val Arg Asp Arg Gly Ser Val Ser Asp Tyr Ile Leu
                165                 170                 175

Ser Glu Ile Lys Pro Leu His Asn Phe Leu Thr Gln Leu Gln Pro Phe
            180                 185                 190

Leu Lys Leu Lys Gln Lys Gln Ala Asn Leu Val Leu Lys Ile Ile Glu
        195                 200                 205

Gln Leu Pro Ser Ala Lys Glu Ser Pro Asp Lys Phe Leu Glu Val Cys
    210                 215                 220

Thr Trp Val Asp Gln Ile Ala Ala Leu Asn Asp Ser Lys Thr Arg Lys
225                 230                 235                 240
```

s1546PCT2.ST25

Thr Thr Ser Glu Thr Val Arg Ala Val Leu Asp Ser Leu Ser Glu Lys
245 250 255

Lys Lys Ser Ser Pro Ala Ala Asp
260

<210> 2

<211> 795

<212> DNA

<213> artificial sequence

<220>

<223> single chain I-Cre I

<400> 2
atggccaaca ctaagtacaa taaagaattt ctcctgtatc tggcaggttt cgtcgacggc 60

gatggctcca ttatcgcaca gatcaagccg aatcagagct acaagtttaa acaccaactg 120

tctctcactt tccaggttac ccagaaaact caacgtcgct ggttcctgga taagctggta 180

gatgagatcg gtgtgggcta tgtacgcgac cgtggctctg tgagcgacta tatcctgtct 240

gagattaaac cactgcataa ttttctgacc cagctgcagg ctatgctgga gcgtatccgt 300

ctgttcaaca tgcgtgagtt cctgctgtac ctggccggct ttgtggacgg tgacggtagc 360

atcatcgctc agattaaacc aaaccagtct tataaattca agcatcagct gtccctgacc 420

tttcaggtga ctcaaaagac ccagcgccgt tggtttctgg acaaactggt ggatgaaatt 480

ggcgttggtt acgtacgtga tcgcggtagc gtttccgatt acattctgag cgaaatcaag 540

ccgctgcaca acttcctgac tcaactgcaa ccgtttctga aactgaaaca gaaacaggca 600

aacctggttc tgaaaattat cgaacagctg ccgtctgcaa aagaatcccc ggacaaattc 660

ctggaagttt gtacctgggt ggatcagatt gcagctctga acgattctaa gacgcgtaaa 720

accacttctg aaaccgttcg tgctgtgctg gacagcctga gcgagaagaa gaaatcctcc 780

ccggcggccg actag 795

<210> 3

<211> 492

<212> DNA

<213> artificial sequence

<220>

<223> natural I-Cre I

s1546PCT2.ST25

<400> 3
atggccaata ccaaatataa caaagagttc ctgctgtacc tggccggctt tgtggacggt 60
gacggtagca tcatcgctca gattaaacca aaccagtctt ataaattcaa gcatcagctg 120
tccctgacct ttcaggtgac tcaaaagacc cagcgccgtt ggtttctgga caaactggtg 180
gatgaaattg gcgttggtta cgtacgtgat cgcggtagcg tttccgatta cattctgagc 240
gaaatcaagc cgctgcacaa cttcctgact caactgcaac cgtttctgaa actgaaacag 300
aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg 360
gacaaattcc tggaagtttg tacctgggtg gatcagattg cagctctgaa cgattctaag 420
acgcgtaaaa ccacttctga aaccgttcgt gctgtgctgg acagcctgag cgagaagaag 480
aaatcctccc cg 492

<210> 4
<211> 492
<212> DNA
<213> artificial sequence

<220>
<223> non homologous I-Cre I

<400> 4
atggccaaca ctaagtacaa taaagaattt ctcctgtatc tggcaggttt cgtcgacggc 60
gatggctcca ttatcgcaca gatcaagccg aatcagagct acaagtttaa acaccaactg 120
tctctcactt tccaggttac ccagaaaact caacgtcgct ggttcctgga taagctggta 180
gatgagatcg gtgtgggcta tgtacgcgac cgtggctctg tgagcgacta tatcctgtct 240
gagattaaac cactgcataa ttttctgacc cagctgcagc cgttcctcaa gctgaagcaa 300
aaacaggcca atctcgtgct gaagatcatt gagcaactgc catccgccaa agagtctccg 360
gataaatttc tggaggtctg cacttgggtt gaccaaatcg ctgcactcaa cgactccaaa 420
acccgcaaga cgaccagcga gactgtacgc gcagttctgg attctctctc cgaaaaaaag 480
aagtctagcc cg 492

<210> 5
<211> 492
<212> DNA
<213> artificial sequence

<220>
<223> template I-Cre I

s1546PCT2.ST25

<400> 5
atggccaata ccaaatataa caaagagttc ctgctgtacc tggccggctt tgtggacggt 60

gacggtagca tcatcgctca gattaaacca aaccagtctt ataagtttaa acatcagcta 120

agcttgacct ttcaggtgac tcaaaagacc cagcgccgtt ggtttctgga caaactagtg 180

gatgaaattg gcgttggtta cgtacgtgat cgcggatccg tttccaacta catcttaagc 240

gaaatcaagc cgctgcacaa cttcctgact caactgcagc cgtttctgaa actgaaacag 300

aaacaggcaa acctggttct gaaaattatc gaacagctgc cgtctgcaaa agaatccccg 360

gacaaattcc tggaagtttg tacctgggtg gatcagattg cagctctgaa cgattctaag 420

acgcgtaaaa ccacttctga aaccgttcgt gctgtgctgg acagcctgag cgagaagaag 480

aaatcctccc cg 492

<210> 6

<211> 42

<212> DNA

<213> artificial sequence

<220>

<223> UlibfIor primer

<400> 6
acgacggcca gtgaattcac catggccaat accaaatata ac 42

<210> 7

<211> 75

<212> DNA

<213> artificial sequence

<220>

<223> UlibIrev primer

<400> 7
cacctgaaag gtcaagctta gmbbatgttt aaacttmbba gactgmbbtg gmbbaatmbb 60

agcgatgatg ctacc 75

<210> 8

<211> 49

<212> DNA

<213> artificial sequence

<220>

<223> UlibIIfor primer

<400> 8
gtttaaacat cagctaagct tgacctttvv kgtgactcaa aagacccag 49

<210> 9

<211> 45

<212> DNA

<213> artificial sequence

<220>

<223> UlibIIrev primer

<400> 9
gatgtagttg gaaacggatc cmbbatcmbb tacgtaacca acgcc 45