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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 February 2013 (07.02.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/019745 A1

(51) International Patent Classification:
A61K 38/00 (2006.01) *A61K 31/7088* (2006.01)

(21) International Application Number:
PCT/US2012/048884

(22) International Filing Date:
30 July 2012 (30.07.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/513,491 29 July 2011 (29.07.2011) US

(71) Applicant *(for all designated States except US)*: **GEORGIA HEALTH SCIENCES UNIVERSITY** [US/US]; 1120 15th Street CJ-3301, Augusta, GA 30912 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **MEILER, Steffen** [US/US]; 787 Sparkleberry Road, Evans, GA 30809 (US). **DYNAN, William** [US/US]; 428 Wateroak Lane, Martinez, GA 30907 (US).

(74) Agents: **PABST, Patrea, L.** et al.; Pabst Patent Group LLP, 1545 Peachtree Street, N.E., Suite 320, Atlanta, GA 30309 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/019745 A1

(54) **Title**: METHODS AND COMPOSITIONS FOR GENETICALLY MODIFIYING CELLS

(57) **Abstract**: Methods and compositions for genetically modifying cells via a receptor mediated pathway are provided. The methods and compositions be used to treat blood disorders such as sickle cell disease. Preferred receptors include transferrin receptor and the CXCR4 receptors.

METHODS AND COMPOSITIONS FOR GENETICALLY MODIFIYING CELLS

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was made with Government Support under Grant No. PN2EY018244 awarded by the National Institutes of Health. The Government has certain rights in the invention.

FIELD OF THE INVENTION

The invention is generally related to compositions and methods for genetically modifying cells. In particular, the methods and compositions are based on receptor mediated delivery.

BACKGROUND OF THE INVENTION

Several diseases result from problems at the genetic level. Common genetic based diseases include sickle cell disease, cystic fibrosis, and certain cancers. Current methods of treating these genetic based diseases are often limited to treating the symptoms of the disease. Although technology exists for correcting genetic defects at the molecular level, translation of this technology into effective methods of gene therapy has been elusive.

Sickle cell disease is caused by a point mutation in the β-globin chain of hemoglobin. Therefore, gene therapy for sickle cell disease should target cells that express or will develop to express mutated β-globin and should avoid cells that do not express mutated β-globin. Selecting which cells to genetically modify will determine whether the gene therapy will ultimately be successful.

Zinc finger nucleases (ZFN) have been used to target specific areas of a gene and then cleave the DNA at the location of the gene. Custom-designed ZFNs that combine the non-specific cleavage domain of Fok I endonuclease with zinc finger binding domains offer a general way to deliver site-specific double-strand breaks to the genome, and stimulate local homologous recombination by several orders of magnitude. This makes targeted gene correction or genome editing a viable option in human cells.

By using template DNA that contains the correct DNA sequence, repair of the DNA after nuclease cleavage can eliminate the genetic defect. One problem with the ZFN approach is that it can lead to unwanted DNA damage if not targeted to the specific site of interest and if left in the cell for extended periods.

It is therefore an object of the invention to provide compositions and methods for repairing DNA in cells including somatic cells, embryonic cells, stem cells, progenitor cells, pluripotent cells, or multipotent cells.

It is still another object of the invention to provide compositions and methods for treating genetic disorders.

SUMMARY OF THE INVENTION

Methods and compositions for genetically modifying cells, preferably multipotent cells, are provided. Exemplary methods include obtaining cells to be modified from a subject having a genetic disorder to be treated, culturing the cells, contacting the cultured cells with a repair nucleic acid, contacting the cultured cells with a conjugate containing a fusion protein capable of selectively promoting homologous recombination of the repair nucleic acid into the cellular DNA to correct the genetic disorder, and transplanting the modified cells back into the subject. The fusion protein is conjugated to a ligand for a cell surface receptor expressed on the cultured cells. In one embodiment, the cell surface protein or receptor that is expressed is an endogenous cell surface protein or receptor. In other embodiments, the receptor is a heterologous receptor encoded by a transgene. The cargo-ligand conjugate binds the expressed cell surface protein via the ligand and the conjugate is endocytosed into the cell.

An exemplary endogenous cell surface protein includes, but is not limited to the transferrin receptor (TfR) or a chemokine receptor. A preferred chemokine receptor is CXCR4. Thus, ligands for TfR or CXCR4 can be conjugated to the fusion protein to promote endocytosis of the conjugate into the cell via the ligand-receptor interaction.

In one embodiment, the fusion protein is a zinc finger nuclease, for example a zinc finger nuclease that targets the β-globin gene.

The compositions can be used to genetically modify cells. The genetic modifications include genomic modifications to replace mutated or non-functional genes or segments of a gene. In one embodiment, the zinc finger nuclease cleaves specifically within a target gene, for example within the beta globin gene. The zinc finger nuclease protein in combination with a repair nucleic acid sequence to be inserted within the target gene is delivered to the cell. The zinc finger nuclease induces a double-stranded break or a nick at a preselected site near the mutation in the target gene. The broken DNA ends will enter the homologous recombination pathway which will incorporate the repair nucleic acid sequence.

The genetically modified cells can be used to treat genetic disorders including blood disorders such as sickle cell disease. Blood cell progenitor cells can be cultured under conditions that stimulate expression of cell surface proteins or receptors including but not limited to TfR. The contacted cells are harvested and administered to a subject.

In one embodiment, the methods can also involve administering a template DNA strand optionally in combination with an inhibitor of nonhomologous end joining recombination. The template DNA can contain a desired nucleic acid sequence that encodes a functional protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic diagram showing the principle of the transferrin receptor mediated delivery system.

Figure 2 shows a cartoon drawing of purified GFP-ZFN bound to GFP. A schematic drawing of the linearized pEGFP-N1 plasmid is also shown. Also provided is an agarose gel analysis of the cleaved GFP plasmid.

Figure 3 shows the temperature, pH, and electrolyte-dependence of in vitro cleavage activity of purified ZFNs.

Figure 4 shows agarose gel analyses of in vitro cleavage activity of transferrin-ZFNs. Both concentration- and time-dependent results are provided.

Figure 5 shows a flow chart of the zinc finger nuclease and transferrin-zinc finger nuclease conjugate preparation as well as an SDS-PAGE analysis of the purified zinc finger nucleases.

Figure 6A, 6B and 6C shows a schematic diagram of a zinc finger nuclease (A), a diagram of the synthesis of the transferrin-zinc finger nuclease conjugate (B) and SDS-PAGE analyses of the conjugate (C).

Figures 7A and 7B show the results of an in vitro cutting assay of a GFP target sequence. A range of concentrations of the zinc finger nuclease and transferrin-zinc finger nucleases were analyzed.

Figures 8A, 8B and 8C show time dependent expression of TfR in human hematopoietic stem and progenitor cells. A) A bar graph of cell expansion (fold increase) versus days. The inset shows a line graph of cell number versus days. B) A bar graph of flow cytometry data showing % gated cells versus days. The inset is a bar graph of mean fluorescent intensity (MFI) versus days. C) Flow cytometry data showing the cell surface expression levels of Thy-1, CD34 and TfR.

Figure 9 shows the time-dependent expression of the TfR in cultured mouse hematopoietic stem and progenitor cells.

Figures 10A and 10B are bar graphs showing time- and dose-dependent nuclear uptake of tf-ZFN in 293 cells, respectively.

Figures 11A and 11B show flow cytometric analysis of human CD34+ hematopoietic stem cells revealed that >99% of human CD34+ HSPCs internalized the fusion protein after incubation with the conjugate.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of skill in the art to which the disclosed method and compositions belong. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present method and compositions, the particularly useful methods, devices, and materials are as described. Publications cited herein and the material for which they are cited are hereby specifically incorporated by reference. Nothing herein is to be

construed as an admission that the present invention is not entitled to antedate such disclosure by virtue of prior invention. No admission is made that any reference constitutes prior art. The discussion of references states what their authors assert, and applicants reserve the right to challenge the accuracy and pertinency of the cited documents.

The singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a moiety" includes a plurality of such moieties, reference to "the moiety" is a reference to one or more moieties and equivalents thereof known to those skilled in the art, and so forth.

As used herein, the terms "bind," "binding," "interact," and "interacting," refer to covalent interactions, noncovalent interactions and steric interactions. A covalent interaction is a chemical linkage between two atoms or radicals formed by the sharing of a pair of electrons (a single bond), two pairs of electrons (a double bond) or three pairs of electrons (a triple bond). Covalent interactions are also known in the art as electron pair interactions or electron pair bonds. Noncovalent interactions include, but are not limited to, van der Waals interactions, hydrogen bonds, weak chemical bonds (via short-range noncovalent forces), hydrophobic interactions, ionic bonds and the like. A review of noncovalent interactions can be found in Alberts et al., in Molecular Biology of the Cell, 3d edition, Garland Publishing, 1994. Steric interactions are generally understood to include those where the structure of the compound is such that it is capable of occupying a site by virtue of its three dimensional structure, as opposed to any attractive forces between the compound and the site.

Cleavage or cleaving of nucleic acids refers to the breakage of the covalent backbone of a nucleic acid molecule. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible, and double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. DNA cleavage can result in the production of either blunt ends or staggered ends.

In certain embodiments, fusion polypeptides are used for targeted double-stranded DNA cleavage.

The terms "nucleic acid," "polynucleotide," and "oligonucleotide" are used interchangeably and refer to a deoxyribonucleotide or ribonucleotide polymer, in linear or circular conformation, and in either single- or double-stranded form. For the purposes of the present disclosure, these terms are not to be construed as limiting with respect to the length of a polymer. The terms can encompass known analogues of natural nucleotides, as well as nucleotides that are modified in the base, sugar and/or phosphate moieties (e.g., phosphorothioate backbones). In general, an analogue of a particular nucleotide has the same base-pairing specificity; i.e., an analogue of A will base-pair with T.

The terms "polypeptide," "peptide" and "protein" are used interchangeably to refer to a polymer of amino acid residues. The term also applies to amino acid polymers in which one or more amino acids are chemical analogues or modified derivatives of a corresponding naturally-occurring amino acids.

"Recombination" refers to a process of exchange of genetic information between two polynucleotides. As used herein, "homologous recombination (HR)" refers to the specialized form of such exchange that takes place, for example, during repair of double-strand breaks in cells. This process requires nucleotide sequence homology, uses a template nucleic acid molecule to repair a "target" molecule (i.e., the one that experienced the double-strand break). HR leads to the transfer of genetic information from the template, or donor, to the target. HR often results in an alteration of the sequence of the target molecule such that part or all of the sequence of the template is incorporated into the target polynucleotide.

As used herein, "subject" includes, but is not limited to, a vertebrate, more specifically a mammal (e.g., a human, horse, pig, rabbit, dog, sheep, goat, non-human primate, cow, cat, guinea pig or rodent), a fish, a bird, a reptile or an amphibian. The subject may be an invertebrate, more specifically an arthropod (e.g., insects and crustaceans). The term does not

denote a particular age or sex. Thus, adult and newborn subjects, as well as fetuses, whether male or female, are intended to be covered.

A "zinc finger DNA binding protein" (or binding domain) is a protein, or a domain within a larger protein, that binds DNA in a sequence-specific manner through one or more zinc fingers. The term zinc finger DNA binding protein is often abbreviated as zinc finger protein or ZFP.

II. Compositions for Genetic Modifications

Compositions for genetically modifying cells to include a fusion protein conjugated or linked to a ligand for a cell surface receptor protein. The fusion protein is separated from the ligand after the composition is internalized by the cell. Typically, the composition is internalized by endocytosis. Representative fusion proteins include a DNA binding domain and a nuclease domain. The fusion proteins can be produced using conventional molecular biology techniques.

In a preferred embodiment, a fusion protein conjugate includes a DNA binding domain and a Type IIS nuclease domain. The fusion protein can be releasably conjugated to transferrin or stromal cell-derived factor 1. The DNA binding domain of the fusion protein can include a zinc finger. The DNA binding domain of the fusion protein can include a transcription activator-like (TAL) effector. The Type IIS nuclease can be *Fok*I.

A. DNA Binding Domains

Preferred DNA binding domains include zinc fingers and transcription activator-like (TAL) effectors. Other DNA binding domains include HMG domains, helix-turn-helix, homeodomains, leucine zipper, helix-loop-helix, and histone folds. The disclosed fusion proteins can have any combination of DNA binding domains. Preferably, the DNA binding domains recognize and bind to specific nucleic acid sequences in a target gene. Representative fusion proteins include zinc finger nucleases.

1. HMG Domain

Generally, the HMG domain includes a global fold of three helices stabilized in an `L-shaped' configuration by two hydrophobic cores. The high mobility group chromosomal proteins HMG1 or HMG2, which are common to all eukaryotes, bind DNA in a non-sequence-specific fashion, for

example to promote chromatin function and gene regulation. They can interact directly with nucleosomes and are believed to be modulators of chromatin structure. They are also important in activating a number of regulators of gene expression, including p53, Hox transcription factors and steroid hormone receptors, by increasing their affinity for DNA. HMG proteins include HMG-1/2, HMG-I(Y) and HMG-14/17.

The HMG-1/2-box proteins can be further distinguished into three subfamilies according to the number of HMG domains present in the protein, their specific of sequence recognition and their evolutionary relationship. The first group contains chromosomal proteins bound to DNA with no sequence specificity (class I, HMG1 and HMG2), the second contains ribosomal and mitochondrial transcription factors which show sequence specificity in the presence of another associating factor when bound with DNA (class II, yeast ARS binding protein ABF-2, UBF and mitochondrial transcription factor mtTF-1), and the third contains gene-specific transcription factors which show sequence specific DNA binding (class III, lymphoid enhancer-binding factors LEF-1 and TCF-1; the mammalian sex-determining factor SRY, and the closely related SOX proteins; and the fungal regulatory proteins Mat-MC, Mat-a1, Ste11 and Rox1). The HMG1/2-box DNA binding domain is about 75 to about 80 amino acids and contains highly conserved proline, aromatic and basic residues. Common properties of HMG domain proteins include interaction with the minor groove of the DNA helix, binding to irregular DNA structure, and the capacity to modulate DNA structure by bending.

SOX (SRY-type HMG box) proteins have critical functions in a number of developmental processes, including sex determination, skeleton formation, pre-B and T cell development and neural induction. SOX9 plays a direct role during chondrogenesis by binding and activating the chondrocyte-spacific enhancer of the Col2a1 gene. Loss of SOX9 gene function leads to the genetic condition known as Campomelic Dysplsia (CD), a form of dwarfism characterized by extreme skeletal malformation, and one in which three-quarters of XY individual are either intersexes or exhibit male to female sex reversal. There are more than 20 members cloned in SOX family.

All of which contain an HMG domain, which can bind specifically to the double strand DNA motif and shares >50% identify with the HMG domain of SRY, the human testis-determining factor. The preferred DNA-binding site of SOX9 have been defined to be AGAACAATGG (SEQ ID NO: 1), which contains the SOX core-binding element (SCBE), AACAAT, flanking 5' AG and 3' GG nucleotides enhance binding by SOX9.

In one embodiment, the DNA binding domain has at least one HMG box domain, generally at least two, more particularly 2-5 HMG box domains. The HMG box domain can bind to an AT rich DNA sequence, for example, using a large surface on the concave face of the protein, to bind the minor groove of the DNA. This binding bends the DNA helix axis away from the site of contact. The first and second helices contact the DNA, their N-termini fitting into the minor groove whereas helix 3 is primarily exposed to solvent. Partial intercalation of aliphatic and aromatic residues in helix 2 occurs in the minor groove.

In other embodiments, the fusion protein can have at least one polynucleotide binding domain, typically two or more polynucleotide binding domains. The polynucleotide binding domains can be the same or different. For example, the polynucleotide-binding polypeptide can include at least one HMG box in combination with one or more DNA binding domains selected from the group consisting of an HMG box, homeodomain and POU domain; zinc finger domain such as C_2H_2 and C_2C_2; amphipathic helix domain such as leucine zipper and helix-loop-helix domains; and histone folds. The polynucleotide binding domain can be specific for a specific polynucleotide sequence, or preferably non-specifically binds to a polynucleotide. Alternatively, the polynucleotide-binding polypeptide can have a combination of at least one polynucleotide binding domain that binds in a sequence specific manner and at least one polynucleotide binding-domain that binds DNA non-specifically.

2. Helix-turn-helix

Certain embodiments provide DNA binding domains having a helix-turn-helix motif or at least a polynucleotide binding region of a helix-turn-helix protein. Helix-turn-helix proteins have a similar structure to bacterial

regulatory proteins such as the lambda repressor and cro proteins, the lac repressor and so on which bind as dimers and their binding sites are palindromic. They contain 3 a helical regions separated by short turns which is why they are called helix-turn-helix proteins. One protein helix (helix 3) in each subunit of the dimer occupies the major groove of two successive turns of the DNA helix. Thus, in another embodiment, the disclosed DNA-binding polypeptides can form dimers or other multi-component complexes, and have 1 to 3 helices.

3. Homeodomain

In yet another embodiment, the DNA binding domain includes a homeodomain or a portion of a homeodomain protein. Homeodomain proteins bind to a sequence of 180 base pairs initially identified in a group of genes called homeotic genes. Accordingly, the sequence was called the homeobox. The 180 bp corresponds to 60 amino acids in the corresponding protein. This protein domain is called the homeodomain. Homeodomain-containing proteins have since been identified in a wide range of organisms including vertebrates and plants. The homeodomain shows a high degree of sequence conservation. The homeodomain contains 4 α helical regions. Helices II and III are connected by 3 amino acids comprising a turn. This region has a very similar structure to helices II and III of bacterial DNA binding proteins.

4. Zinc Finger

Yet another embodiment provides a DNA binding domain having a zinc finger domain or at least a portion of a zinc finger protein. Zinc finger proteins have a domain with the general structure: Phe (sometimes Tyr) - Cys - 2 to 4 amino acids - Cys - 3 amino acids - Phe (sometimes Tyr) - 5 amino acids - Leu - 2 amino acids - His - 3 amino acids - His. The phenylalanine or tyrosine residues which occur at invariant positions are required for DNA binding. Similar sequences have been found in a range of other DNA binding proteins though the number of fingers varies. For example, the SP1 transcription factor which binds to the GC box found in the promoter proximal region of a number of genes has 3 fingers. This type of

zinc finger which has 2 cysteines and 2 histidines is called a C_2H_2 zinc finger.

Another type of zinc finger which binds zinc between 2 pairs of cysteines has been found in a range of DNA binding proteins. The general structure of this type of zinc finger is: Cys - 2 amino acids - Cys - 13 amino acids - Cys - 2 amino acids - Cys. This is called a C_2C_2 zinc finger. It is found in a group of proteins known as the steroid receptor superfamily, each of which has 2 C_2C_2 zinc fingers.

The DNA-binding domain of a ZFN may be composed of two to six zinc fingers. Each zinc finger motif is typically considered to recognize and bind to a three-base pair sequence and as such, a protein including more zinc fingers targets a longer sequence and therefore has a greater specificity and affinity to the target site.

Zinc finger binding domains can be "engineered" to bind to a predetermined nucleotide sequence. See, for example, Beerli et al. (2002) Nature Biotechnol. 20:135-141; Pabo et al. (2001) Ann. Rev. Biochem. 70:313-340; Isalan et al. (2001) Nature Biotechnol. 19:656-660; Segal et al. (2001) Curr. Opin. Biotechnol. 12:632-637; Choo et al. (2000) Curr. Opin. Struct. Biol. 10:411-416. Consequently, zinc finger binding domains can be engineered to have a novel binding specificity, compared to a naturally-occurring zinc finger protein. Engineering methods include, but are not limited to, rational design and various types of empirical selection methods. Rational design includes, for example, using databases comprising triplet (or quadruplet) nucleotide sequences and individual zinc finger amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of zinc fingers which bind the particular triplet or quadruplet sequence. See, for example, U.S. Pat. Nos. 6,140,081; 6,453,242; 6,534,261; 6,610,512; 6,746,838; 6,866,997; 7,067,617; U.S. Patent Application Publication Nos. 2002/0165356; 2004/0197892; 2007/0154989; 2007/0213269; and International Patent Application Publication Nos. WO 98/53059 and WO 2003/016496 all of which are incorporate by reference in their entireties.

In one embodiment, the zinc finger binding domain binds specifically to a sequence of the β-globin gene. Other genes of interest for targeting in blood disorders are well known in the art, such as the F9 gene for treating hemophilia.

5. Leucine zipper

Another embodiment provides a DNA binding domain having a leucine zipper or at least a portion of a leucine zipper protein. The first leucine zipper protein was identified from extracts of liver cells, and it was called C/EBP because it is an enhancer binding protein and it was originally thought to bind to the CAAT promoter proximal sequence. C/EBP will only bind to DNA as a dimer. The region of the protein where the two monomers join to make the dimer is called the dimerization domain. This lies towards the C-terminal end of the protein. When the amino acid sequence was examined it was found that a leucine residue occurs every seventh amino acid over a stretch of 35 amino acids. If this region were to form an a helix then all of these leucines would align on one face of the helix.

Because leucine has a hydrophobic side chain, one face of the helix is very hydrophobic. The opposite face has amino acids with charged side chains which are hydrophilic. The combination of hydrophobic and hydrophilic characteristics gives the molecule is amphipathic moniker. Adjacent to the leucine zipper region is a region of 20-30 amino acids which is rich in the basic (positively charged) amino acids lysine and arginine. This is the DNA binding domain - often referred to as the bZIP domain - the basic region of the leucine zipper. C/EBP is thought to bind to DNA by these bZIP regions wrapping round the DNA helix

The leucine zipper - bZIP structure has been found in a range of other proteins including the products of the *jun* and *fos* oncogenes. Whereas C/EBP binds to DNA as a homodimer of identical subunits, fos cannot form homodimers at all and jun/jun homodimers tend to be unstable. However fos/jun heterodimers are much more stable. These fos/jun heterodimers correspond to a general transcription factor called AP1 which binds to a variety of promoters and enhancers and activates transcription. The consensus AP1 binding site is TGACTCA which is palindromic.

6. Helix-loop-helix

Another embodiment provides a DNA binding domain having helix-loop-helix domain or a polynucleotide binding portion of a helix-loop-helix protein. Helix-loop-helix proteins are similar to leucine zippers in that they form dimers via amphipathic helices. They were first discovered as a class of proteins when a region of similarity was noticed between two enhancer binding proteins called E47 and E12. This conserved region has the potential to form two amphipathic separated by a loop hence helix-loop-helix. Next to the dimerization domain is a DNA binding domain, again rich in basic amino acids and referred to as the bHLH domain. These structures are also found in a number of genes required for development of the Drosophila nervous system - the Achaete-scute complex, and in a protein called MyoD which is required for mammalian muscle differentiation.

7. Histone Fold

In still another embodiment, the DNA binding domain includes a histone polypeptide, a fragment of a histone polypeptide, or at least one histone fold. Histone folds exist in histone polypeptides monomers assembled into dimers. Histone polypeptides include H2A, H2B, H3, and H4 which can form heterodimers H2A-2B and H3-H4. It will be appreciated that histone-like polypeptides can also be used in the disclosed compositions and methods. Histone-like polypeptides include, but are not limited to, HMf or the histone from *Methanothermous fervidus*, other archaeal histones known in the art, and histone-fold containing polypeptides such as MJ1647, CBF, TAFII or transcription factor IID, SPT3, and Dr1-DRAP (Sanderman, K., et al., *Cell. Mol. Life Sci.* 54:1350-1364 (1998), which is incorporated by reference in its entirety).

8. TALENS

The DNA binding domain can include transcription activator-like (TAL) effectors. Produced by plant pathogenic bacteria in the genus *Xanthomonas*, the native function of these proteins is to directly modulate host gene expression. Upon delivery into host cells via the bacterial type III secretion system, TAL effectors enter the nucleus, bind to effector-specific sequences in host gene promoters and activate transcription. Their targeting

specificity is determined by a central domain of tandem, 33–35 amino acid repeats, followed by a single truncated repeat of 20 amino acids. Methods of engineering TAL to bind to specific nucleic acids are described in Cermak, et al., Nucl. Acids Res. (2011) 1-11. US Patent Publication 20110145940, which is incorporated by reference in its entirety discloses TAL effectors and methods of using them to modify DNA.

B. Nuclease Domain

The nuclease domain preferably includes Fok I nuclease. Other enzymes or portions of enzymes that modify DNA can also be used. For example, proteins having nucleic acid modifying domains include, but are not limited to, transferases (e.g., terminal deoxynucleotidyl transferase), RNases (RNase A, ribonuclease H), DNases (DNase I), ligases (T4 DNA ligase, E. coli DNA ligase), nucleases (Fok I nuclease), kinases (T4 polynucleotide kinase), phoshatases (calf intestinal alkaline phosphatase, bacterial alkaline phosphatase), exonucleases (X exonuclease), endonucleases, glycosylases (uracil DNA glycosylases), deaminases (activation-induced deaminase, AID) and the like. These proteins can target specific nucleic acid sequences.

The nuclease may be naturally or non-naturally occurring. The nuclease domain portion of the fusion proteins can be obtained from an endonuclease or exonuclease. Exemplary endonucleases from which a nuclease domain can be derived include, but are not limited to, restriction endonucleases and homing endonucleases. See, for example, 2002-2003 Catalogue, New England Biolabs, Beverly, Mass.; and Belfort et al. (1997) Nucleic Acids Res. 25:3379-3388. Additional enzymes which cleave DNA are known (e.g., Si Nuclease; mung bean nuclease; pancreatic DNase I; micrococcal nuclease; yeast HO endonuclease; see also Linn et al. (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993). One or more of these enzymes (or functional fragments thereof) can be used as a source of nuclease domains and half-domains. Restriction endonucleases (restriction enzymes) are present in many species and are capable of sequence-specific binding to DNA (at a recognition site), and cleaving DNA at or near the site of binding.

Certain restriction enzymes (e.g., Type IIS) cleave DNA at sites removed from the recognition site and have separable binding and cleavage domains. For example, the Type IIS enzyme Fok I catalyzes double-stranded cleavage of DNA, at 9 nucleotides from its recognition site on one strand and 13 nucleotides from its recognition site on the other. See, for example, U.S. Pat. Nos. 5,356,802; 5,436,150 and 5,487,994; as well as Li et al. (1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. USA 90:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim et al. (1994b) J. Biol. Chem. 269:31,978-31,982. Thus, in one embodiment, the fusion protein include a nuclease domain from at least one Type IIS restriction enzyme.

An exemplary Type IIS restriction enzyme, whose cleavage domain is separable from the binding domain, is Fok I. This particular enzyme is active as a dimer. Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA 95: 10,570-10,575.

The non-specific nuclease domain of Fok I is functionally independent of its natural DNA-binding domain. The domain must dimerize to accomplish a double-strand break.

C. Nuclear Localization Signal

The fusion protein optionally includes a nuclear localization signal. Nuclear localization signals (NLS) or domains are known in the art and include for example, SV 40 T antigen or a fragment thereof, such as PKKKRKV (SEQ ID NO:2). The NLS can be simple cationic sequences of about 4 to about 8 amino acids, or can be bipartite having two interdependent positively charged clusters separated by a mutation resistant linker region of about 10-12 amino acids. Additional representative NLS include but are not limited to GKKRSKV (SEQ ID NO:3); KSRKRKL (SEQ ID NO:4); KRPAATKKAGQAKKKKLDK (SEQ ID NO:5); RKKRKTEEESPLKDKAKKSK (SEQ ID NO:6); KDCVMNKHHRNRCQYCRLQR (SEQ ID NO:7); PAAKRVKLD (SEQ ID NO:8); and KKYENVVIKRSPRKRGRPRK (SEQ ID NO:9).

D. Zinc Finger Nucleases

An exemplary fusion protein is a zinc finger nucleases (ZFNs). ZFNs are able to introduce double-strand breaks (breaks at the same or very close points in both strands of a double-stranded DNA molecule) at specific locations within a DNA molecule which may subsequently be used to disable a specific allele or rewrite the code it contains.

In order to form a dimer, two ZFN molecules must meet with their respective recognition sites not less than 4-6 base pairs apart but also not so far apart that they may not dimerize. While one ZFN molecule binds its target sequence on one strand, another ZFN molecule binds its target sequence on the opposite strand. The nuclease domains dimerize and each cleaves its own strand, producing a double stranded break.

Methods of making and testing ZFN to evaluate their activity in vivo are known in the art (Maeder, Mol. Cell., 31(2):294-301). Photoactivatable ZFN can be turned on and off by light of different wavelengths. The photoactivatable ZFNs can be produced by introducing an azobenzene derivative spanning two strategically located surface cysteine residues. In one embodiment, the cysteine residues are in or near the α5 helix of the Fok I nuclease domain. The distance between the two engineered cysteine residues is ~8Å, which matches the cis-configuration of azobenzene. Illumination with 465nm blue light or dark adaptation can isomerize the azobenzene to the thermodynamically more stable trans-configuration and disrupt the α5 helix. This alteration can prevent dimerization of the nuclease domain which leads to turning off its DNA cleavage activity.

The OPEN zinc finger protein selection method can be used to generate zinc fingers that recognize a sequence that is approximately 30 base pairs from codon 6 of the beta globin gene, the site of the hemoglobin S mutation (Maeder, Mol. Cell., 31(2):294-301). The zinc finger was joined to the natural FokI nuclease domain (See U.S. Patent Publication No. 20110086015 which is incorporated by reference in its entirety). In certain embodiments the natural FokI nuclease domain can be replaced with a codon optimized nuclease domain Beta-globin ZFNs containing the natural FokI domain were 60-70% as active as our standard GFP ZFNs. By contrast, those

containing the codon optimized domain were 200-300% as active as the GFP-ZFNs.

A representative ZFN (referred to as GFPZFN1) is encoded by the following nucleic acid sequence

ATGGACTACAAAGACGATGACGACCCAAAAAAGAAGCGAAAGGT
ACCCTTCCAGTGTCGCATTTGCATGCGGAACTTTTCGACGCGGCA
GAAGCTTGGCGTGCATACCCGTACTCATACCGGTGAAAAACCGTT
TCAGTGTCGGATCTGTATGCGAAATTTCTCCGTCGCGCACAACTTG
ACCCGCCATCTACGTACGCACACCGGCGAGAAGCCATTCCAATGC
CGAATATGCATGCGCAACTTCAGTCAGCACCCCAACCTGACGAGG
CACCTAAAAACCCACACAGGACAGAAGGACCAACTAGTCAAAAG
TGAACTGGAGGAGAAGAAATCTGAACTTCGTCATAAATTGAAATA
TGTGCCTCATGAATATATTGAATTAATTGAAATTGCCAGAAATTC
CACTCAGGATAGAATTCTTGAAATGAAGGTAATGGAATTTTTTAT
GAAAGTTTATGGATATAGAGGTAAACATTTGGGTGGATCAAGGA
AACCGGACGGAGCAATTTATACTGTCGGATCTCCTATTGATTACG
GTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAATCTGC
CAATTGGCCAAGCAGATGAAATGCAACGATATGTCGAAGAAAAT
CAAACACGAAACAAACATATCAACCCTAATGAATGGTGGAAAGT
CTATCCATCTTCTGTAACGGAATTTAAGTTTTTATTTGTGAGTGGT
CACTTTAAAGGAAACTACAAAGCTCAGCTTACACGATTAAATCAT
ATCACTAATTGTAATGGAGCTGTTCTTAGTGTAGAAgAGCTTTTAA
TTGGtGGAGAAATGATTAAAGCCGGCACATTAAcCTTAGAGGAAG
TGAGAcgGAAATTTAATAACGGCGAGatAAACTTTTAATCTAGA
(SEQ ID NO:10)

or having an amino acid sequence of

MDYKDDDDPKKKRKVPFQCRICMRNFSTRQKLGVHTRTHTGEKPFQ
CRICMRNFSVAHNLTRHLRTHTGEKPFQCRICMRNFSQHPNLTRHLK
THTGQKDQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRIL
EMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKA
YSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFK
FLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTL
TLEEVRRKFNNGEINF
(SEQ ID NO:11)

Another ZFN (also referred to as GFPZFN2) is encoded by

ATGGACTACAAAGACGATGACGACCCAAAAAAGAAGCGAAAGGT
ACCCTTCCAGTGTCGCATTTGCATGCGGAACTTTTCGGCGCCGAG
CAAGCTTGACAGGCATACCCGTACTCATACTGGTGAAAAACCGTT
TCAGTGTCGGATCTGTATGCGAAATTTCTCCGATCGCTCTAATCTG
ACCCGCCATCTACGTACGCACACCGGCGAGAAGCCATTCCAATGC
CGAATATGCATGCGCAACTTCAGTGAGGGGGGGAACCTGATGAG
GCACCTAAAAACCCACACAGGACAGAAGGACCAACTAGTCAAAA
GTGAACTGGAGGAGAAGAAATCTGAACTTCGTCATAAATTGAAAT
ATGTGCCTCATGAATATATTGAATTAATTGAAATTGCCAGAAATT

CCACTCAGGATAGAATTCTTGAAATGAAGGTAATGGAATTTTTTA
TGAAAGTTTATGGATATAGAGGTAAACATTTGGGTGGATCAAGGA
AACCGGACGGAGCAATTTATACTGTCGGATCTCCTATTGATTACG
GTGTGATCGTGGATACTAAAGCTTATAGCGGAGGTTATAATCTGC
CAATTGGCCAAGCAGATGAAATGCAACGATATGTCGAAGAAAAT
CAAACACGAAACAAACATATCAACCCTAATGAATGGTGGAAAGT
CTATCCATCTTCTGTAACGGAATTTAAGTTTTTATTTGTGAGTGGT
CACTTTAAAGGAAACTACAAAGCTCAGCTTACACGATTAAATCAT
ATCACTAATTGTAATGgAGCTGTTCTTAGTGTAGAAGAGCTTTTAA
TTGGTGGAGAAATGATTAAAGCCGGCACATTAACCTTAGAGGAA
GTGAGAcGGAAATTTAATAACGGCGAGaTAAACTTT
(SEQ ID NO:12)

or has the following amino acid sequence

MDYKDDDDPKKKRKVPFQCRICMRNFSAPSKLDRHTRTHTGEKPFQ
CRICMRNFSDRSNLTRHLRTHTGEKPFQCRICMRNFSEGGNLMRHLK
THTGQKDQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRIL
EMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKA
YSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFK
FLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTL
TLEEVRRKFNNGEINF
(SEQ ID NO:13)

It will be appreciated that the N-terminal methionine can be removed.

GFP-ZFN1(SP202B) recognizes Target Site: 5' GAA GAT GGT 3' (bp 300-292 of GFP)

GFP-ZFN2 (SP202A) recognizes Target Site: 5' GAC GAC GGC 3' (bp 307-315 of GFP)

GFPZFN1 and GFPZFN2 recognize and cleave the gene encoding green fluorescent protein.

D. Cell Surface Receptor Ligands

The nuclease fusion proteins are conjugated to ligands that specifically recognize receptors expressed on the surface of the cells to be genetically modified. Preferred ligands specifically bind to cell surface receptors expressed by stem cells or progenitor cells. These receptors include, but are not limited to transferrin and chemokine receptors.

1. Transferrin

A Transferrin receptor ligand (TfR ligand) includes transferrin but can be any molecule that specifically binds to the Transferrin receptor (TfR). TfR ligands bind to TfR and are internalized and processed via an endosomal

pathway. Upon binding of the TfR ligand to TfR, the receptor and the ligand (bound to the cargo) are endocytosed in clathrin coated endosomal vesicles. The clathrin coated vesicles transport the TfR ligand and TfR to the endosome. Within the endosome, cleavage of the disulfide bond between the TfR ligand and the cargo occurs. The cargo is released into the cytosol while the TfR ligand and TfR are recycled to the cell surface. At the cell surface, the neutral pH allows dissociation of the TfR ligand from the TfR.

TfR are expressed on all nucleated cells in the body. Rapidly dividing cells express high levels of TfR. For example, tumor cells or cultured cell lines can have 10,000 to 100,000 molecules per cell (Inoue *et al.* J Cell Physiol 156:212-217, 1993). Nonproliferating cells have low or undetectable levels of TfR.

Transferrin (Tf) binds to the transferrin receptor (TfR) on the surface of cells. The term transferrin encompasses transferrin and isolated peptide fragments that specifically bind to the TfR. Transferrin can be commercially obtained as holo-transferrin or apo-transferrin, or can be recombinantly-produced or chemically synthesized using any method established in the art. Transferrin variants can be used.

In one embodiment, TfR monoclonal antibodies (mAbs), Tf oligomers and Gambogic acid can be used as TfR ligands (Lim and Shen, Pharm Res. 2004; 21(11): 1985-92; Yazdi et al., Cancer Res. 1995; 55(17):3763-71; Pandey et al. Blood 110:3517-3525, 2007). Cytotoxin conjugates used to target TfR have been explored (Yazdi et al., Cancer Res. 1995; 55(17):3763-71; Yazdi and Murphy, Cancer Res. 1994; 54(24):6387-94; Wenning et al., Biotechnol Bioeng. 1998; 57(4):484-96).

2. Chemokines

The chemokine ligand can be a ligand that binds to one of the following chemokine receptors CCR1, CCR2, CCR3, CCR4, CCR5, CCR6, CCR7, CCR8, CCR9, CCR10, CXCR1, CXCR2, CXCR3, CXCR4, CXCR5, CXCR6, CXCR7, CX3CR1, and XCR1. Preferred ligands bind to one of the following receptors CXCR1, CXCR4, CXCR5, CXCR6, CCR6, CCR8, CCR9, CCR10, XCR1 and CX3CR1. Even more preferred is chemokine

receptor CXCR4. The ligand for CXCR4 includes stromal cell-derived factor-1 alpha (SDF-1α) and stromal cell-derived factor-1 beta (SDF-1β).

CXCR4 is a chemokine receptor in the GPCR gene family, and is expressed by cells in the immune system and the central nervous system. In response to binding its ligand SDF-1, CXCR4 triggers the migration and recruitment of immune cells. In addition to acting as a chemokine receptor, CXCR4 is a co-receptor for entry of HIV into T cells.

CXCR4 induces downstream signaling by several different pathways. As a GPCR, CXCR4 binding of SDF-1 activates G-protein mediated signaling, including downstream pathways such as ras, and PI3 kinase. PI3 kinase activated by SDF-1 and CXCR4 plays a role in lymphocyte chemotaxis in response to these signals. One endpoint of CXCR4 signaling is the activation of transcription factors such as AP-1 and chemokine regulated genes. JAK/STAT signaling pathways also appear to play a role in SDF-1/CXCR4 signaling.

E. Methods of Making the Conjugates

The fusion protein and the ligand for the cell surface receptor can be associated covalently or non-covalently, directly or indirectly, or with or without a linker.

A covalent link, or like terms, refers to an intermolecular association or bond which involves the sharing of electrons in the bonding orbitals of two atoms. A non-covalent link, or like terms, refers to intermolecular interaction among two or more separate molecules or molecular entities which does not involve a covalent bond. Intermolecular interaction is dependent upon a variety of factors, including, for example, the polarity of the involved molecules, and the charge (positive or negative), if any, of the involved molecules. Non-covalent associations are selected from ionic interactions, dipole-dipole interactions, van der Waal's forces, and combinations thereof.

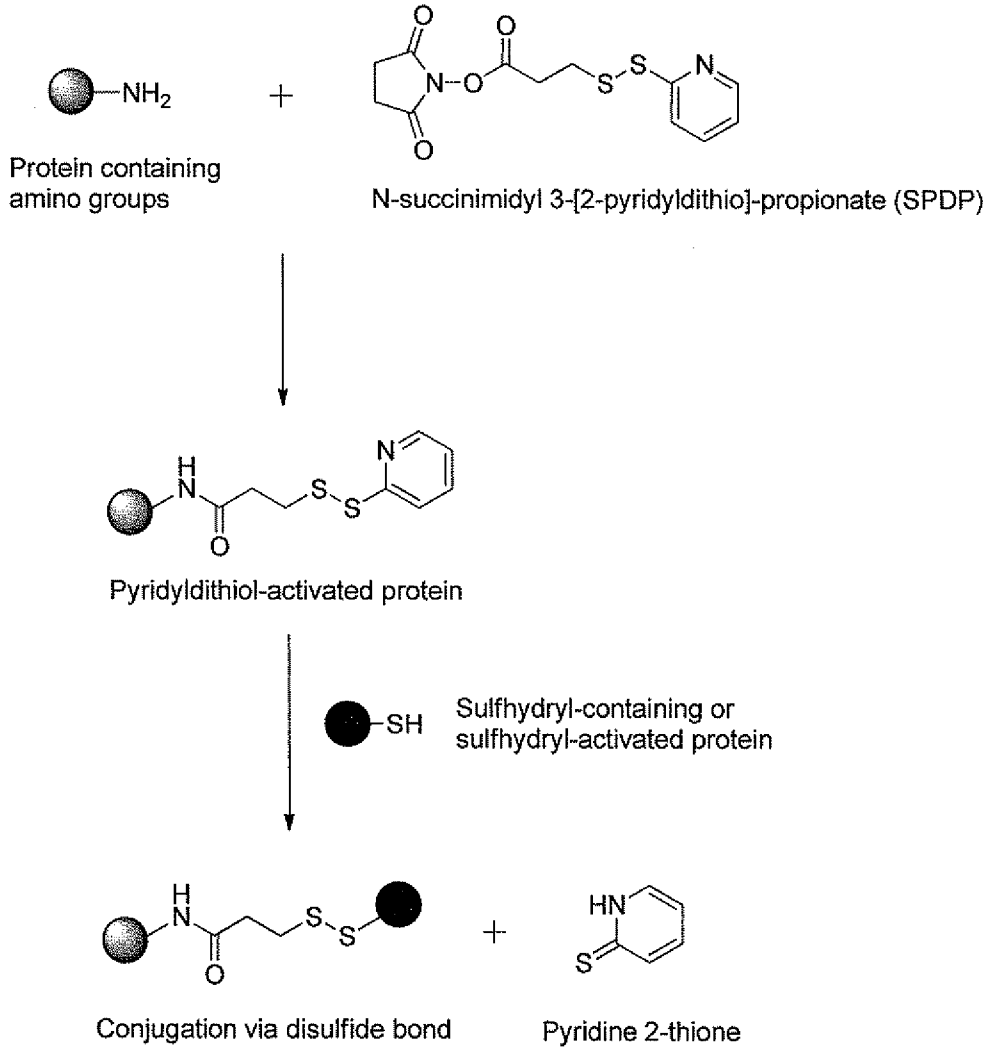
Protein crosslinkers that can be used to crosslink the cargo composition to the disclosed peptide are known in the art and are defined based on utility and structure and include DSS (Disuccinimidylsuberate), DSP (Dithiobis(succinimidylpropionate)), DTSSP (3,3'-Dithiobis

(sulfosuccinimidylpropionate)), SULFO BSOCOES (Bis[2-(sulfosuccinimdooxycarbonyloxy) ethyl]sulfone), BSOCOES (Bis[2-(succinimdooxycarbonyloxy)ethyl]sulfone), SULFO DST (Disulfosuccinimdyltartrate), DST (Disuccinimdyltartrate), SULFO EGS (Ethylene glycolbis(succinimidylsuccinate)), EGS (Ethylene glycolbis(sulfosuccinimidylsuccinate)), DPDPB (1,2-Di[3'-(2'-pyridyldithio) propionamido]butane), BSSS (Bis(sulfosuccinimdyl) suberate), SMPB (Succinimdyl-4-(p-maleimidophenyl) butyrate), SULFO SMPB (Sulfosuccinimdyl-4-(p-maleimidophenyl) butyrate), MBS (3-Maleimidobenzoyl-N-hydroxysuccinimide ester), SULFO MBS (3-Maleimidobenzoyl-N-hydroxysulfosuccinimide ester), SIAB (N-Succinimidyl(4-iodoacetyl) aminobenzoate), SULFO SIAB (N-Sulfosuccinimidyl(4-iodoacetyl)aminobenzoate), SMCC (Succinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate), SULFO SMCC (Sulfosuccinimidyl-4-(N-maleimidomethyl)cyclohexane-1-carboxylate), NHS LC SPDP (Succinimidyl-6-[3-(2-pyridyldithio) propionamido) hexanoate), SULFO NHS LC SPDP (Sulfosuccinimidyl-6-[3-(2-pyridyldithio)propionamido)hexanoate), SPDP (N-Succinimidyl-3-(2-pyridyldithio) propionate), NHS BROMOACETATE (N-Hydroxysuccinimidylbromoacetate), NHS IODOACETATE (N-Hydroxysuccinimidyliodoacetate), MPBH (4-(N-Maleimidophenyl) butyric acid hydrazide hydrochloride), MCCH (4-(N-Maleimidomethyl)cyclohexane-1-carboxylic acid hydrazide hydrochloride), MBH (m-Maleimidobenzoic acid hydrazidehydrochloride), SULFO EMCS (N-(epsilon-Maleimidocaproyloxy) sulfosuccinimide), EMCS (N-(epsilon-Maleimidocaproyloxy) succinimide), PMPI (N-(p-Maleimidophenyl) isocyanate), KMUH (N-(kappa-Maleimidoundecanoic acid) hydrazide), LC SMCC (Succinimidyl-4-(N-maleimidomethyl)-cyclohexane-1-carboxy(6-amidocaproate-)), SULFO GMBS (N-(gamma-Maleimidobutryloxy) sulfosuccinimide ester), SMPH (Succinimidyl-6-(beta-maleimidopropionamidohexanoate)), SULFO KMUS (N-(kappa-Maleimidoundecanoyloxy)sulfosuccinimide ester), GMBS (N-(gamma-Maleimidobutyrloxy) succinimide), DMP (Dimethylpimelimidate

hydrochloride), DMS (Dimethylsuberimidate hydrochloride), MHBH(Wood's Reagent) (Methyl-p-hydroxybenzimidate hydrochloride, 98%), DMA (Dimethyladipimidate hydrochloride).

The linkers of the disclosed conjugates can be selected from the group consisting of an amino acid linker, a peptide linker, and an alkyl linker including at least two linked carbon atoms.

In one embodiment, the TfR ligand, Tf, is linked to a ZFN via an amine-sulfhydryl cross-linking as shown in the schematic synthesis below.



The conjugates can be purified to remove or reduce contaminants to obtain a pharmaceutically acceptable composition. The purified composition is typically 85%, 90%, 95%, or 99% free of contaminants.

F. Protein/Peptide Variants

The fusion protein can include a variant polypeptides for the DNA binding domain, the nuclease, or both. Protein variants and derivatives are well understood by those of skill in the art and can involve amino acid sequence modifications. For example, amino acid sequence modifications typically fall into one or more of three classes: substitutional, insertional or deletional variants. Insertions include amino and/or carboxyl terminal fusions as well as intrasequence insertions of single or multiple amino acid residues. Insertions ordinarily will be smaller insertions than those of amino or carboxyl terminal fusions, for example, on the order of one to four residues. Deletions are characterized by the removal of one or more amino acid residues from the protein sequence. Typically, no more than about from 2 to 6 residues are deleted at any one site within the protein molecule. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the protein, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture.

Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis and PCR mutagenesis. Amino acid substitutions are typically of single residues, but can occur at a number of different locations at once; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. Substitutions, deletions, insertions or any combination thereof can be combined to arrive at a final construct. The mutations must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. Substitutional variants are those in which at least one residue has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Table 1 and are referred to as conservative substitutions.

TABLE 1: Amino Acid Substitutions
Original Residue Exemplary Conservative Substitutions, others are known in the art.

Ala	Ser
Arg	Lys; Gln
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn, Lys
Glu	Asp
Gly	Pro
His	Asn;Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the protein properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine, in this case, (e) by increasing the number of sites for sulfation and/or glycosylation.

For example, the replacement of one amino acid residue with another that is biologically and/or chemically similar is known to those skilled in the art as a conservative substitution. For example, a conservative substitution would be replacing one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as, for example, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. Such conservatively substituted variations of each explicitly disclosed sequence are included within the mosaic polypeptides provided herein.

Substitutional or deletional mutagenesis can be employed to insert sites for N-glycosylation (Asn-X-Thr/Ser) or O-glycosylation (Ser or Thr). Deletions of cysteine or other labile residues also may be desirable. Deletions or substitutions of potential proteolysis sites, e.g. Arg, is accomplished for example by deleting one of the basic residues or substituting one by glutaminyl or histidyl residues.

Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and asparyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the o-amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, Proteins: Structure and Molecular Properties, W. H. Freeman & Co., San Francisco pp 79-86 [1983]), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

Specifically disclosed are variants of TfR ligands, zinc finger nucleases, and other polypeptides herein disclosed which have at least, 65%, 70%, 75%, 80%, 85%, 90% or 95% homology to the stated sequence. Those of skill in the art readily understand how to determine the homology of two proteins. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison can be

conducted by the local homology algorithm of Smith and Waterman Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. MoL Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A. 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. Science 244:48-52, 1989, Jaeger et al. Proc. Natl. Acad. Sci. USA 86:7706-7710, 1989, Jaeger et al. Methods Enzymol. 183:281-306, 1989 which are herein incorporated by reference for at least material related to nucleic acid alignment.

It is understood that the description of conservative mutations and homology can be combined together in any combination, such as embodiments that have at least 70% homology to a particular sequence wherein the variants are conservative mutations.

As this specification discusses various proteins and protein sequences it is understood that the nucleic acids that can encode those protein sequences are also disclosed. This would include all degenerate sequences related to a specific protein sequence, i.e. all nucleic acids having a sequence that encodes one particular protein sequence as well as all nucleic acids, including degenerate nucleic acids, encoding the disclosed variants and derivatives of the protein sequences. Thus, while each particular nucleic acid sequence may not be written out herein, it is understood that each and every sequence is in fact disclosed and described herein through the disclosed protein sequence.

It is understood that there are numerous amino acid and peptide analogs which can be incorporated into the disclosed peptides. For example, there are numerous D amino acids or amino acids which have a different functional substituent. The opposite stereo isomers of naturally occurring peptides are disclosed, as well as the stereo isomers of peptide analogs. These amino acids can readily be incorporated into polypeptide chains by

charging tRNA molecules with the amino acid of choice and engineering genetic constructs that utilize, for example, amber codons, to insert the analog amino acid into a peptide chain in a site specific way (Thorson et al., Methods in Molec. Biol. 77:43-73 (1991), Zoller, Current Opinion in Biotechnology, 3:348-354 (1992); Ibba, Biotechnology & Genetic Engineering Reviews 13:197-216 (1995), Cahill et al., TIBS, 14(10):400-403 (1989); Benner, TIB Tech, 12:158-163 (1994); Ibba and Hennecke, Bio/technology, 12:678-682 (1994) all of which are herein incorporated by reference at least for material related to amino acid analogs).

Molecules can be produced that resemble peptides, but which are not connected via a natural peptide linkage. For example, linkages for amino acids or amino acid analogs can include CH_2NH--, --CH_2S--, --CH_2--CH_2 --, --CH=CH-- (cis and trans), --$COCH_2$ --, --$CH(OH)CH_2$--, and --CHH_2SO— (These and others can be found in Spatola, A. F. in Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., Vega Data (March 1983), Vol. 1, Issue 3, Peptide Backbone Modifications (general review); Morley, Trends Pharm Sci (1980) pp. 463-468; Hudson, D. et al., Int J Pept Prot Res 14:177-185 (1979) (--CH_2NH--, CH_2CH_2--); Spatola et al. Life Sci 38:1243-1249 (1986) (--CH H_2--S); Hann J. Chem. Soc Perkin Trans. I 307-314 (1982) (--CH--CH--, cis and trans); Almquist et al. J. Med. Chem. 23:1392-1398 (1980) (--$COCH_2$--); Jennings-White et al. Tetrahedron Lett 23:2533 (1982) (--$COCH_2$--); Szelke et al. European Appln, EP 45665 CA (1982): 97:39405 (1982) (--$CH(OH)CH_2$--); Holladay et al. Tetrahedron. Lett 24:4401-4404 (1983) (--$C(OH)CH_2$--); and Hruby Life Sci 31:189-199 (1982) (--CH_2--S--); each of which is incorporated herein by reference. A particularly preferred non-peptide linkage is --CH_2NH--. It is understood that peptide analogs can have more than one atom between the bond atoms, such as b-alanine, g-aminobutyric acid, and the like.

Amino acid analogs and analogs and peptide analogs often have enhanced or desirable properties, such as, more economical production, greater chemical stability, enhanced pharmacological properties (half-life,

absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity, and others.

D-amino acids can be used to generate more stable peptides, because D amino acids are not recognized by peptidases and such. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (e.g., D-lysine in place of L-lysine) can be used to generate more stable peptides. Cysteine residues can be used to cyclize or attach two or more peptides together. This can be beneficial to constrain peptides into particular conformations. (Rizo and Gierasch Ann. Rev. Biochem. 61:387 (1992), incorporated herein by reference).

G. Pharmaceutical Carriers

The disclosed compositions and conjugates can include a pharmaceutically acceptable carrier. By "pharmaceutically acceptable" is meant a material that is not biologically or otherwise undesirable, i.e., the material can be administered to a subject without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the pharmaceutical composition in which it is contained. The carrier would naturally be selected to minimize any degradation of the active ingredient and to minimize any adverse side effects in the subject, as would be well known to one of skill in the art.

Suitable carriers and their formulations are described in Remington: The Science and Practice of Pharmacy (19th ed.) ed. A. R. Gennaro, Mack Publishing Company, Easton, Pa. 1995. Typically, an appropriate amount of a pharmaceutically acceptable salt is used in the formulation to render the formulation isotonic. Examples of the pharmaceutically acceptable carrier include, but are not limited to, saline, Ringer's solution and dextrose solution. The pH of the solution is preferably from about 5 to about 8, and more preferably from about 7 to about 7.5. Further carriers include sustained release preparations such as semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, liposomes or microparticles. It will be apparent to those persons skilled in the art that certain carriers may be more preferable

depending upon, for instance, the route of administration and concentration of composition being administered.

Pharmaceutical carriers are known to those skilled in the art. These most typically would be standard carriers for administration of drugs to humans or cultured cells, including solutions such as sterile water, saline, and buffered solutions at physiological pH. The compositions can be administered according to standard procedures used by those skilled in the art. For example, pipetting is a common technique for administering compositions to cultured cells.

Pharmaceutical compositions can include carriers, thickeners, diluents, buffers, preservatives, surface active agents and the like in addition to the molecule of choice. Pharmaceutical compositions can also include one or more active ingredients such as antimicrobial agents, anti-inflammatory agents, anesthetics, and the like.

III. Methods of Use

The disclosed fusion proteins and conjugates can be used to deliver specific cargo to cells, in particular, to deliver fusion proteins that can modify the DNA of a cell. The cargo to be delivered can be coupled to ligand for a cell surface receptor. Preferably, the cell surface receptor is induced to be expressed or over-expressed while the cells are being cultured. In a preferred embodiment, the fusion protein is a protein capable of modifying cellular DNA.

The disclosed conjugates are typically added to a culture of cells obtained from a subject to be treated. The cells are optionally expanded in culture prior to contacting the cells with the conjugates. Cell culture conditions can be adjusted to induce cell surface receptors to be expressed. For example, stimulating proliferation of cells can increase the expression of TfR. Culturing techniques for stimulating proliferation of cells are known in the art (Ogawa Blood 81:2844-2853, 1993). Temperature, pH, and electrolyte-dependence can be modulated to increase receptor binding or biological activity.

Delivery of the donor template DNA can be performed using common techniques known to those of skill in the art. For example, the

donor template can be delivered using electroporation, microinjection, or transfection.

In some embodiments, a method for treating a subject having sickle cell disease includes obtaining hematopoietic stem cells from a subject, delivering template DNA into the cytoplasm of the hematopoietic stem cells while in culture, contacting the hematopoietic stem cells containing the template DNA while in culture with a fusion protein conjugated to transferrin or stromal cell-derived factor 1. The fusion protein can include a DNA binding domain specific for the beta globin gene and a nuclease domain that induces a double-stranded break or nick in the beta globin gene of the hematopoietic cells and promote homologous recombination with the template DNA to replace the A to T nucleotide substitution in the beta globin gene of the hematopoietic stem cells. The method can include selecting the cultured hematopoietic cells that have internalized the fusion protein conjugate and undergone homologous recombination with the template DNA and transplanting the hematopoietic cells that have undergone homologous recombination with the template DNA into the subject having sickle cell disease. The transplanted hematopoietic cells typically produce healthy red blood cells in the subject. The transplanted hematopoietic stem cells can be CD34+. The hematopoietic stem cells can be cultured for 2 to 7 days prior to transplantation into the subject. The fusion protein can a zinc finger nuclease, for example, *Fok*I. The DNA binding domain can include a transcription activator-like (TAL) effector.

In some embodiments, a method of treating a genetic disorder in a subject can include a) culturing stem cells or progenitor cells under conditions that increase expression of transferrin receptors; b) delivering to the stem cells or progenitor cells template DNA including wildtype DNA that can replace genomic DNA of stem cells or progenitor cells including a mutation resulting in the genetic disorder; c) contacting the stem cells or progenitor cells containing the template DNA with a zinc finger nuclease conjugated to transferrin wherein the stem cells or progenitor cell containing the template DNA internalize the zinc finger nuclease conjugated to transferrin; and d) selecting the stem cells or progenitor cells that have

undergone homologous recombination with the template DNA and administering these stem cells or progenitor cells to the subject. The genetic disorder can be sickle cell disease. The zinc finger nuclease can target the β-globin gene. The method can include delivering an inhibitor of nonhomologous end joining to the hematopoietic stem cells containing the template DNA.

In some embodiments, a method for genetically modifying genomic DNA of cell can including culturing the cell under conditions that promote expression of a cell surface receptor selected from the group consisting of transferrin receptor and CXCR4; delivering template DNA to be incorporated into the genomic DNA of the cell to the nucleus of the cell; contacting the cell with a fusion protein that specifically binds to the genomic DNA in the cell corresponding to the template DNA and promotes homologous recombination of the template DNA into the genomic DNA of the cell.

A. Methods of Treating Disorders

The disclosed conjugates can be used to treat genetic disorders or diseases including genetic blood disorders such as sickle cell disease. In one embodiment, the treatment of the genetic disorder includes culturing hematopoietic stem cells under conditions that stimulate proliferation, introducing template DNA into the cultured cells, and then contacting the cultured hematopoietic stem cells with a conjugate composition, preferably a zinc finger nuclease conjugated to a transferrin or stromal derived growth factor 1. The zinc finger nuclease specifically cuts a target gene and the template DNA is inserted into the cut via homologous recombination to produce a genetically modified hematopoietic stem cell. The modified hematopoietic stem cells can then be transplanted back into the subject.

The template DNA contains the corrected sequence for the targeted gene. For example, the template DNA can contain the wild-type or non-mutant sequence for the β-globin gene.

In one embodiment, the method of treating a blood disorder also involves the administration of an inhibitor of nonhomologous end joining recombination. As disclosed herein, pushing the cell towards homologous

recombination is important for proper repair of the double stranded nucleic acid break created by the ZFN. Examples of inhibitors of NEJH are disclosed herein.

A preferred embodiment provides a method for treating one or more symptoms of a genetic disorder including obtaining stem cells or progenitor cells from a subject to be treated. Preferably the stem cells are hematopoietic stem cells or hematopoietic progenitor cells. The cells are cultured in vivo, optionally to increase expression of a cell-surface receptor such as the transferrin receptor. Zinc finger nuclease specific for a targeted gene is administered to the cultured cells in combination with a repair nucleic acid to be inserted into the targeted gene to correct a mutation or genetic defect in the targeted gene. For example, if the genetic disorder to be treated is sickle cell disease, the repair nucleic acid corresponds to a region of the beta globin gene spanning the causative A to T mutation that responsible for sickle cell disease. The modified stem cells or progenitor cells are then transplanted back into the subject. Although autologous cells are preferred, cells from other individuals can be used to treat the subject. The disclosed methods advantageously use non-viral compositions for inducing the genetic modifications and avoid using foreign transgenes.

The conjugate and donor template can also be administered consecutively or simultaneously. Consecutive administration refers to separate, individual formulations for each composition. The term consecutive administration refers to administration of one composition and then at least 30 minutes later administering the other composition. The consecutive administration can be at least 30 minutes, 45 minutes, 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, 10 hours, 12 hours, 18 hours, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 7 days, 10 days, 14 days or 30 days from the administration of the first composition.

B. Disorders to Be Treated

Disorders include disorders which can be treated, prevented, or otherwise ameliorated by the administration of a composition disclosed herein, e.g., a receptor ligand-cargo conjugate). Suitable disorders include any genetic based disorder, such as sickle cell disease, cystic fibrosis, and

cancer. Blood disorders can also be treated. Blood disorders include any disorder of the blood, blood-forming organs, or blood cells. The disorder or symptoms can be within the blood, a blood cell, or blood-forming organ or the disorder or symptoms can be somewhere else in the body but caused by an abnormality in the blood, a blood cell or blood-forming organ. The term blood disorder includes nutritional anemias (e.g., iron deficiency anemia, sideropenic dysphasia, Plummer-Vinson syndrome, vitamin B_{12} deficiency anemia, vitamin B_{12} deficiency anemia due to intrinsic factor, pernicious anemia, folate deficiency anemia, and other nutritional anemias), myelodysplastic syndrome, bone marrow failure or anemia resulting from chemotherapy, radiation or other agents or therapies, hemolytic anemias (e.g., anemia due to enzyme disorders, anemia due to phosphate dehydrogenase (G6PD) deficiency, favism, anemia due to disorders of glutathione metabolism, anemia due to disorders of glycolytic enzymes, anemias due to disorders of nucleotide metabolism and anemias due to unspecified enzyme disorder), thalassemia (alpha-thalassemia, beta-thalassemia, delta-thalassemia, thalassemia trait, hereditary persistence of fetal hemoglobin (HPFP), and unspecified thalassemias), sickle cell disorders (sickle cell anemia with crisis, sickle cell anemia without crisis, double heterozygous sickling disorders, sickle cell trait and other sickle cell disorders), hereditary hemolytic anemias (hereditary spherocytosis, hereditary elliptocytosis, other hemaglobinopathies and other specified hereditary hemolytic anemias, such as stomatocyclosis), acquired hemolytic anemia (e.g., drug-induced autoimmune hemolytic anemia, other autoimmune hemolytic anemias, such as warm autoimmune hemolytic anemia, drug-induced non-autoimmune hemolytic anemia, hemolytic-uremic syndrome, and other non-autoimmune hemolytic anemias, such as microangiopathic hemolytic anemia); aplastic anemias (e.g., acquired pure red cell aplasia (erythoblastopenia), other aplastic anemias, such as constitutional aplastic anemia and fanconi anemia, acute posthemorrhagic anemic, and anemias in chronic diseases), coagulation defects (e.g., disseminated intravascular coagulation (difibrination syndrome)), hereditary factor VIII deficiency (hemophilia A), hereditary factor IX deficiency

(Christmas disease), and other coagulation defects such as Von Willebrand's disease, hereditary factor XI deficiency (hemophilia C), purpura (e.g., qualitative platelet defects and Glanzmann's disease), neutropenia, agranulocytosis, functional disorders of polymorphonuclear neutrophils, other disorders of white blood cells (e.g., eosinophilia, leukocytosis, lymophocytosis, lymphopenia, monocytosis, and plasmacyclosis), diseases of the spleen, methemoglobinemia, other diseases of blood and blood forming organs (e.g., familial erythrocytosis, secondary polycythemia, essential thrombocytosis and basophilia), thrombocytopenia, infectious anemia, hypoproliferative or hypoplastic anemias, hemoglobin C, D and E disease, hemoglobin lepore disease, and HbH and HbS diseases, anemias due to blood loss, radiation therapy or chemotherapy, or thrombocytopenias and neutropenias due to radiation therapy or chemotherapy, sideroblastic anemias, myelophthisic anemias, antibody-mediated anemias, and certain diseases involving lymphoreticular tissue and reticulohistiocytic system (e.g., Langerhans' cell hystiocytosis, eosinophilic granuloma, Hand-Schuller-Christian disease, hemophagocytic lymphohistiocytosis, and infection-associated hemophagocytic syndrome).

C. Cells

Cells that can be used with the disclosed compositions and methods include multipotent cells, for example stem cells. The stem cells can be adult stem cells, embryonic stem cells, cord blood stem cells, progenitor cells, bone marrow stem cells, induced pluripotent stem cells, totipotent stem cells or hematopoietic stem cells. Representative human cells are CD34+ cells. In a preferred embodiment, the cells do not express a desired surface protein or receptor at significant levels until the expression of the protein or receptor is induced. For example, hematopoietic stem cells or bone marrow cells do not express significant levels of transferrin receptors unless they are induced to express transferrin receptors.

In other embodiments the cells are somatic cells.

The cells can be autologous or heterologous. In certain embodiments, the cells are genetically engineered to express a specific cell surface protein

or receptor. In preferred embodiments, the cells are induced to express an endogenous cell surface protein or receptor.

Examples

Example 1: Zinc Finger Nucleases Using Transferrin-Mediated Endocytosis

Zinc finger nucleases (ZFNs) are custom-designed DNA binding proteins that produce DNA double-strand breaks (DSBs) at predetermined genomic sites, stimulating homology-directed repair in the presence of donor template by many orders of magnitude over the spontaneous rate. The ability to target specific genes with ZFN technology opens therapeutic opportunities for gene correction and selective gene silencing. Sickle cell anemia (SCA) is an ideal disease target because correction of the single gene β-globin mutation in patient-derived, autologous hematopoietic stem progenitor cells (HSPCs) promises to be curative. One of the barriers to ZFN-based gene correction is the lack of a nonviral delivery system that achieves bulk transport of the nucleases to hard-to-transfect target cells, such as embryonic and HSPCs.

The transferrin receptor pathway was selected on the rationale that all nucleated cells, including HSPCs, must import elemental iron to remain viable under *ex vivo* culture conditions. To test the feasibility of this strategy, this initial work used a ZFN pair targeted against a model GFP transgene. Expression was optimized by pilot scale fermentation in an *Escherichia coli* host-vector system and purification to homogeneity by serial chromatography. Purified ZFNs were conjugated to the iron carrier protein, transferrin (tf), using SPDP, an amine and sulfhydryl reactive heterobifunctional crosslinker (Figure 6B). The resulting disulfide linkage is designed to undergo scission ("self-immolation") upon entry into the intracellular reducing environment. *In vitro* DNA cleavage assays and surface plasmon resonance binding assays demonstrated that ZFNs remained competent for target sequence cleavage following conjugation, with only mild to quantitative impairment of activity.

To analyze delivery in biological systems, time- and dose-dependence of tf-mediated ZFN uptake in human osteosarcoma (U2OS 2-6-3) cells was measured. ZFNs in DAPI stained cell nuclei were detected by indirect immunofluorescence and signal intensity was measured in projections of deconvolved depth coded z-stacks. Nuclear uptake of tf-ZFN protein occurred in >95% of cells, was dose-dependent and linear with time in the lower dose ranges, and reached saturation as early as 60 min. Importantly, maximal nuclear uptake was indistinguishable from ZFN plasmid treated cells. These results indicate that endocytic delivery of ZFNs readily traverses the cellular membrane, overcomes the potential hurdle of endosomal trapping, and targets the nucleus with high efficiency.

To demonstrate gene targeting activity, the U2OS 2-6-3 cell assay was used which bears a tandem transgene array at a single locus that is cleavable by the GFP ZFNs. Cells were transfected with lacI-ECFP to mark the target locus, incubated with tf-ZFNs, fixed, and stained for 53BP1, a signaling protein that marks DSBs. Recruitment of 53BP1 to the target locus was observed in 13% (18/135) of tf-ZFN treated cells, whereas no recruitment (0/152) was observed in untreated cells (Table 2). These findings demonstrate that tf-conjugated ZFNs retain cleavage activity after nuclear uptake in a significant percentage of cells.

Table 2

Groups	Cells analyzed	LacI- 53BP1 Co-localization	% Co-localization
Control	152	0	0%
TF-ZFNs	135	18	13.33%

To determine whether the tf-ZFNs are capable of stimulating gene correction, primary mouse adult fibroblasts carrying a mutant GFP transgene were transfected with donor template, incubated with tf-ZFNs, and evaluated cells at 72 h for gene correction as evidenced by GFP expression. Flow cytometry revealed a gene correction rate of 1-2%, identical to ZFN plasmid transfected cells, demonstrating that the technology of shuttling ZFN

proteins to the cell interior via the tf-receptor pathway can deliver bioactive ZFNs to the nuclear compartment, target specific gene sequences, and induce homology-directed repair in the presence of donor DNA.

Example 2: Conjugate Uptake in Progenitor Cells

Mouse (KSL) hematopoietic stem cells were incubated with 120 nM tf-ZFN for 120 min. After uptake, the ligand resides in endosomes in the cytoplasmic compartment. In contrast, tf-ZFN escape the endosome and translocate to the cell nucleus as seen by fluorescence microscopy (data not shown).

Human CD34+ cells were incubated with 120 nM tf-ZFN for 120 min. After uptake, the ligand resides in endosomes, in the cytoplasmic compartment. In contrast, tf-ZFN (sections F-K) escape the endosome and translocate to the cell nucleus as shown by fluorescence microscopy (data not shown). Flow cytometry detected uptake of tf-ZFN in >99% of CD34+ cells, demonstrating the efficacy of this delivery method (Figures 11A and 11B).

Time- and dose-dependent nuclear uptake of tf-ZFN was determined in 293 cells. tf- ZFNs in DAPI stained cell nuclei were detected by indirect immunofluorescence and signal intensity was measured in raw image files of individual z-sections. Nuclear uptake of tf-ZFN protein occurred in >95% of cells, reached saturation at 60 min (Fig. 10A), and quickly tapered off at later time-points. Nuclear uptake of Tf-ZFNs was dose-dependent and exhibited a linear increase of nuclear signal in the lower dose ranges (Fig. 10B). Importantly, maximal nuclear uptake was indistinguishable from ZFN plasmid treated cells. These results confirm that tf-conjugated ZFNs readily traverse the cellular membrane, overcome the potential hurdle of endosomal trapping, and

target the cell nucleus with high efficiency.

Example 3: Receptor Mediated Delivery of ZNF nucleases

Figure 1 shows the general concept of a receptor-mediated delivery system. Recombinant ZFNs are expressed in *E. coli* and conjugated to transferrin via a scissile disulfide bond. These engage the high-affinity transferrin receptor and are taken up into the endosomal compartment by receptor-mediated endocytosis. The disulfide bond is cleaved in the

intracellular reducing environment. The transferrin recycles to the cell surface. The ZFNs escape the endosome and enter the nucleus. It could not have been predicted that the ZFNs would so readily escape the endosome. Endosomal trapping is an issue in many studies of protein uptake. A nuclear localization sequence facilitates entry to the nucleus.

Highly purified transferrin- ZFN conjugates were prepared using model ZFNs (targeted to GFP). Transferrin conjugation was demonstrated to have only a minimal effect on target site binding and enzyme activity in vitro (about a 2-fold loss of activity). ZFN uptake and entry to nucleus were demonstrated in mouse fibroblasts and human cancer cells. The cleavage of GFP target gene array was demonstrated in human cancer cells. Functional correction of mutant GFP was shown in mouse fibroblasts. The presence of TfR in mouse and human hematopoietic stem and progenitor cells following brief expansion in culture was shown. Functionality and uptake kinetics of the TfR pathway was demonstrated in human hematopoietic stem and progenitor cells following brief expansion in culture. Uptake of transferrin-model protein and ZFN conjugates was demonstrated in human and mouse hematopoietic stem and progenitor cells, respectively.

The ability to routinely purify "restriction enzyme grade" ZFN protein was also investigated. Soluble yield is greatly improved by culture using a fermentor (versus shake flask). Two workflows were used – one to purify ZFNS as shown in the left branch of the diagram in Fig. 5, the other to prepare transferrin-ZFN conjugates as shown in the right branch. SDS-PAGE analysis shows the purified ZFNs. The yield is 1-1.5 mg/L culture.

SDS-PAGE under nonreducing conditions shows the conjugates (which do not enter the gel) and under reducing conditions (beta mercapto ethanol) shows cleavage of the scissile (self-immolating) linker to produce free transferrin and ZFN (Figure 6A-C). Fig. 6C also shows an SDS-PAGE analysis wherein the transferrin was tagged with Texas red, so the same or similar gel can be imaged by fluorescence.

Figures 7A and 7B shows that transferrin conjugated ZFNs retain DNA cleavage activity. To demonstrate that the chemistry of the transferrin conjugation reaction did not inactivate the nuclease, the cleavage activity of

wildtype and transferrin-conjugated ZFN proteins was compared in an in vitro cutting assay of a GFP target sequence. The results demonstrate that Tf-conjugated ZFNs require a two-fold higher concentration to achieve wildtype cleavage efficiency.

In another experiment, transferrin was tagged with Texas red and the ZFN is tagged with biotin. Transferrin accumulation is time dependent and little if any transferrin enters the nucleus. The data was obtained from deconvoluted images shown as projections – the small amount of nuclear staining probably arises from sections above or below the nucleus.

Permeabilized cells were fixed and stained with fluorescent streptavidin. Nuclear uptake occurred after 1 hour. Anti-hemagglutinin antibody was also used, and an independent staining method, that shows nuclear uptake (the antibody staining tends to be a little more diffuse so the images are not as crisp, but the result is the same) (data not shown). A more complete time course with the antibody staining indicates that the peak uptake is in between the two times shown here – about 1.5 to 2 hours.

Tf-ZFNs cleave a specific genomic site in a human cancer cell line. The U2OS 2-6-3 cell line was used which bears a gene array that is cleavable by the Tf-conjugated ZFNs. The cells were incubated with a pair of Tf-ZFNs, fixed and processed immunocytochemically for two different markers. LacI-ECFP marks the site of the cleavable gene array in the cell nucleus. 53BP1 is a checkpoint protein that marks double-strand DNA breaks.. Tf-ZFN induced DNA cleavage was defined by the co-localization of LacI and 53BP1 expression.. Evidence for site-specific DNA breaks was found in 13% (18 out of 135 cells) of Tf-ZFN treated cells whereas zero co-localization events were observed in untreated cells. Note the background 53BP1 foci which are characteristic for cancer cell lines. Two representative examples of marker co-localization are shown here (cells 1 and 2).

Tf-ZFNs can correct a single gene mutation in primary mouse fibroblasts (MAFs). MAFs bearing a non-expressing, mutant GFP gene were transfected with a GFP repair DNA template and incubated with a pair of GFP-targeting Tf-ZFNs. Gene correction was monitored by the appearance of GFP+ cells in culture and by indirect immunocyochemistry with an anti-

GFP antibody. Gene correction efficiency was measured by flow cytometry in 1-2% of treated cells which is comparable to viral based delivery methods.

Time-dependent expression of the transferrin receptor in cultured human HSPCs was also investigated. Commercially supplied bone marrow derived human CD34+ HSPCs were cultured for four days under defined growth conditions. Daily cell counts were used to calculate the fold-expansion of CD34+ cells compared to the starting culture. Flow cytometric measurements revealed that 82% of CD34+ cells express the transferrin receptor by day 2 in culture. Additional analysis included the surface marker Thy-1 (CD90) which predicts longterm engraftment potential and showed that 71.5% of cells were triple-positive (CD34, Thy-1, Tf receptor). These results indicate that gene correction using Tf-conjugated ZFNs can be accomplished within two days of the HSPC harvest from the patient when "stemness" is only minimally compromised (Figures 8A-C).

Uptake experiments of labeled transferrin in cultured human HSPCs were also performed. To demonstrate that human HSPCs express a functional transferrin receptor pathway, CD34+ cells were incubated after two days in culture with labeled transferrin and monitored uptake at different time points by deconvolution microscopy. Rapid binding of the labeled ligand to the cell membrane followed by intracellular uptake in endosomal vesicles and recycling of transferrin to the cell surface was observed. The kinetics of the uptake process are consistent with the known biology of the transferrin – transferrin receptor pathway in other cell systems. Flow cytometric analysis revealed that virtually the entire population of Tf receptor expressing CD34+ cells internalized labeled transferrin ligand.

The time-dependent expression of the transferrin receptor in cultured mouse HSPCs was also investigated. Fresh bone marrow derived mouse HSPCs were expanded for seven days in culture under defined growth conditions. Flow cytometry revealed that 70.8% of cells sorted for the stem cell markers c-kit and sca-1 also expressed the transferrin receptor (Figure 9). These results are very similar to those in human HSPCs.

Uptake experiments of a transferrin-conjugated model protein in human HSPCs was investigated. Human CD34+ cells were incubated with a

Tf-Ova protein complex after two days in culture and imaged by deconvolution microscopy. The results convincingly demonstrate that CD34+ cells internalize transferrin conjugated ova-albumin and deposit it in the cell cytoplasm whereas transferrin-free cargo does not gain access to the cell interior.

Mouse HSPCs were expanded in culture for 7 days, lin-depleted and sorted for the surface markers sca-1 and c-kit. Cells were incubated with transferrin-ZFN (100 nM) for 2 h, formalin-fixed, and processed with an anti-FLAG antibody to visualize uptake of the zinc finger nuclease. Single and projections of multiple z-sections clearly show uptake of transferrin-ZFN into the cytoplasm and DAPI-stained nucleus of mouse HSPCs.

We claim:

1. A method for treating a subject having sickle cell disease comprising:

obtaining hematopoietic stem cells from the subject,

delivering template DNA into the cytoplasm of the hematopoietic stem cells while in culture,

contacting the hematopoietic stem cells containing the template DNA while in culture with a fusion protein conjugated to transferrin or stromal cell-derived factor 1, wherein the fusion protein comprises a DNA binding domain specific for the beta globin gene and a nuclease domain that induces a double-stranded break or nick in the beta globin gene of the hematopoietic cells and promotes homologous recombination with the template DNA to replace the A to T nucleotide substitution in the beta globin gene of the hematopoietic stem cells,

selecting the cultured hematopoietic cells that have internalized the fusion protein conjugate and undergone homologous recombination with the template DNA and transplanting the hematopoietic cells that have undergone homologous recombination with the template DNA into the subject having sickle cell disease, wherein the transplanted hematopoietic cells produce healthy red blood cells in the subject.

2. The method of claim 1, wherein the transplanted hematopoietic stem cells are CD34+.

3. The method of claim 1, wherein the hematopoietic stem cells are cultured for 2 to 7 days prior to transplantation into the subject.

4. The method of claim 1, wherein the fusion protein is a zinc finger nuclease.

5. The method of claim 1, wherein the nuclease is *Fok*I.

6. The method of claim 1, wherein the DNA binding domain comprises a transcription activator-like (TAL) effector.

7. A method of treating a genetic disorder in a subject comprising

a) culturing stem cells or progenitor cells under conditions that increase expression of transferrin receptors;

b) delivering to the stem cells or progenitor cells template DNA comprising wildtype DNA that can replace genomic DNA of stem cells or progenitor cells comprising a mutation resulting in the genetic disorder;

c) contacting the stem cells or progenitor cells containing the template DNA with a zinc finger nuclease conjugated to transferrin wherein the stem cells or progenitor cell containing the template DNA internalize the zinc finger nuclease conjugated to transferrin; and

d) selecting the stem cells or progenitor cells that have undergone homologous recombination with the template DNA and administering these stem cells or progenitor cells to the subject.

8. The method of claim 7, wherein the genetic disorder is sickle cell disease.

9. The method of claim 7, wherein the zinc finger nuclease targets the β-globin gene.

10. The method of claim 6, further comprising delivering an inhibitor of nonhomologous end joining to the hematopoietic stem cells containing the template DNA.

11. A fusion protein conjugate comprising a DNA binding domain and a Type IIS nuclease domain wherein the fusion protein is releasably conjugated to transferrin or stromal cell-derived factor 1.

12. The fusion protein of claim 11, wherein the DNA binding domain comprises a zinc finger.

13. The fusion protein of claim 11, wherein the DNA binding domain comprises a transcription activator-like (TAL) effector.

14. The fusion protein of claim 11, wherein the Type IIS nuclease is *Fok*I.

15. A method for genetically modifying genomic DNA of cell comprising

culturing the cell under conditions that promote expression of a cell surface receptor selected from the group consisting of transferrin receptor and CXCR4;

delivering template DNA to be incorporated into the genomic DNA of the cell to the nucleus of the cell;

contacting the cell with a fusion protein according to claim 11, wherein the fusion specifically binds to the genomic DNA in the cell corresponding to the template DNA and promotes homologous recombination of the template DNA into the genomic DNA of the cell.

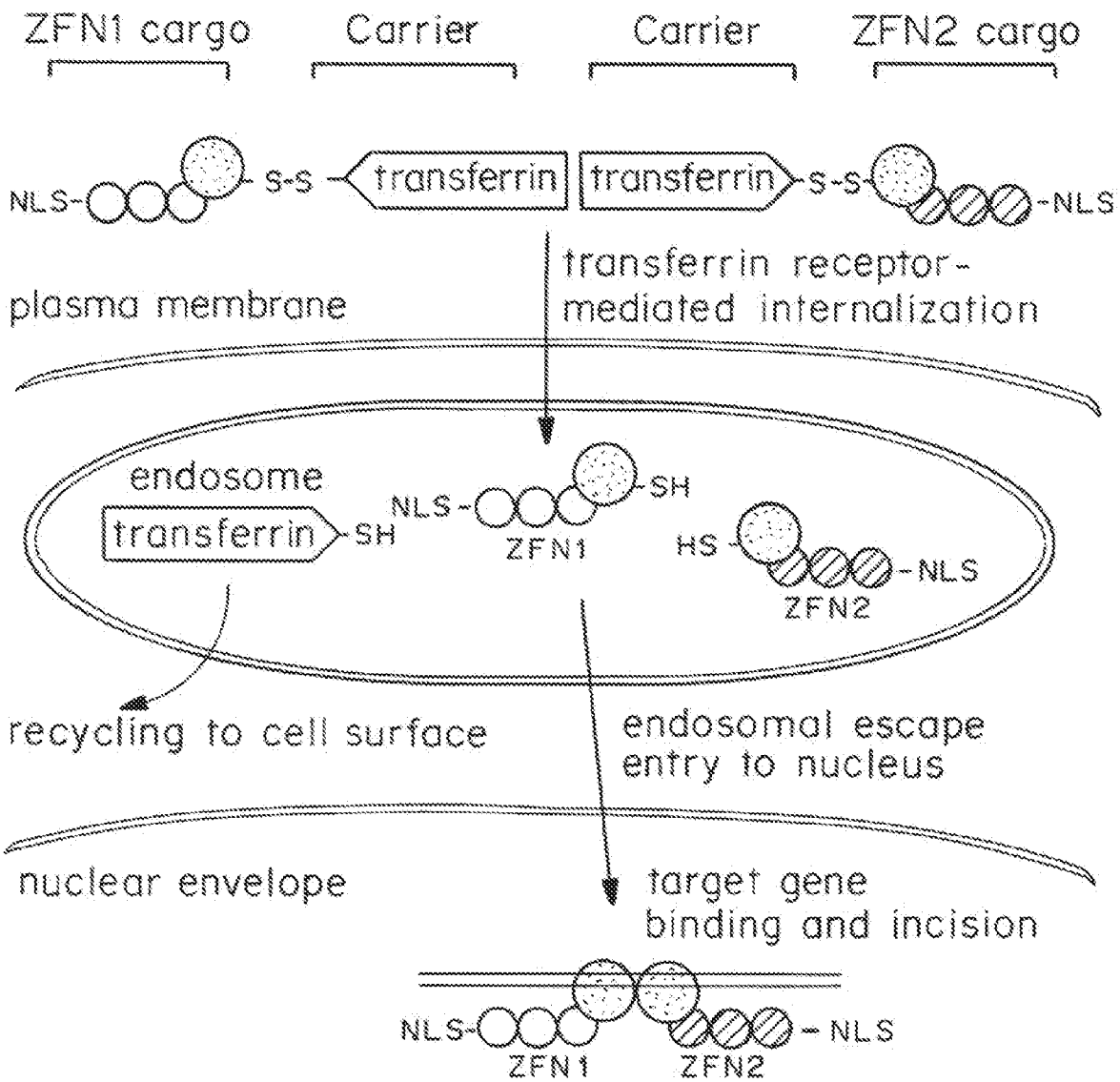

ZFN1 cargo
Carrier
Carrier
ZFN2 cargo
NLS-
s-s
transferrin
transferrin
s-s
-NLS
plasma membrane
transferrin receptor-mediated internalization
endosome
transferrin
SH
NLS-
SH
ZFN1
HS
-NLS
ZFN2
recycling to cell surface
endosomal escape
entry to nucleus
nuclear envelope
target gene
binding and incision
NLS-
-NLS
ZFN1
ZFN2

FIG. 1

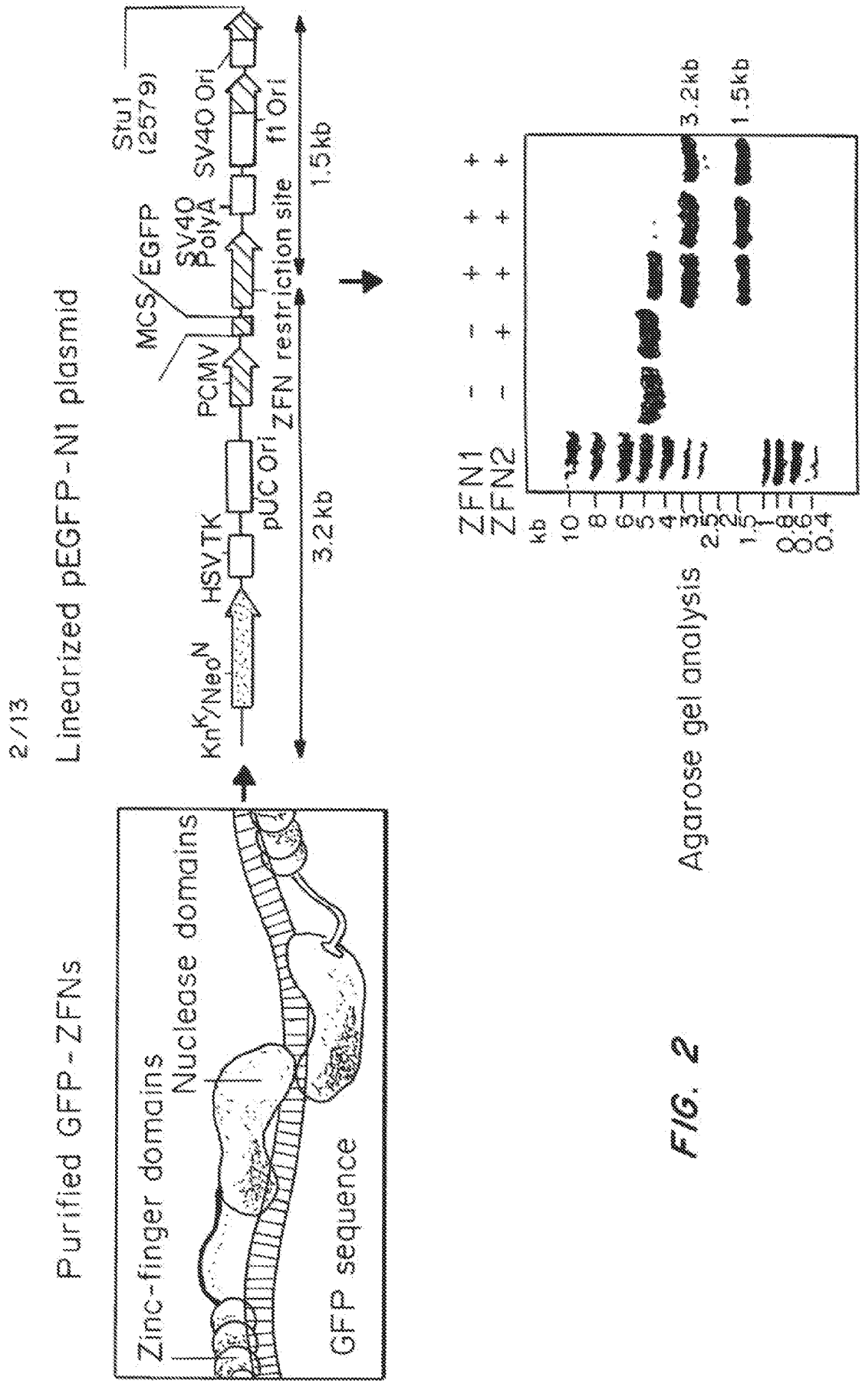

2/13
Purified GFP-ZFNs
Zinc-finger domains
Nuclease domains
GFP sequence
Linearized pEGFP-N1 plasmid
Kn^K/Neo^N
HSV TK
pUC Ori
PCMV
MCS/EGFP
SV40 PolyA
SV40 Ori
f1 Ori
Stu1 (2579)
ZFN restriction site
3.2kb
1.5kb
ZFN1
ZFN2
kb
10
8
6
5
4
3
2.5
2
1.5
1
0.8
0.6
0.4
3.2kb
1.5kb
Agarose gel analysis

FIG. 2

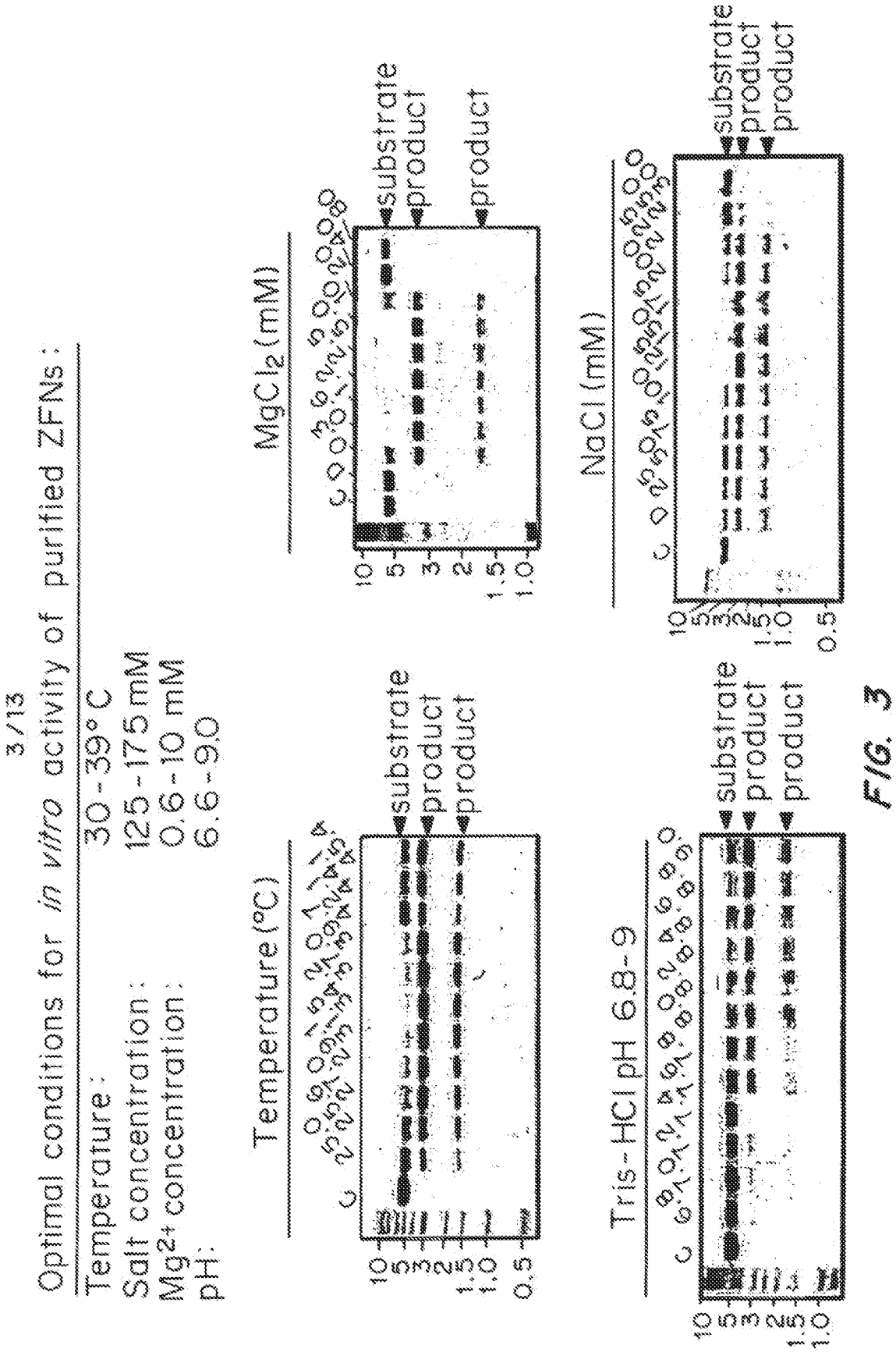

3/13
Optimal conditions for in vitro activity of purified ZFNs:
Temperature: 30-39°C
Salt concentration: 125-175 mM
Mg2+ concentration: 0.6-10 mM
pH: 6.6-9.0
Temperature (°C)
MgCl2 (mM)
NaCl (mM)
Tris-HCl pH 6.8-9
substrate
product
product

FIG. 3

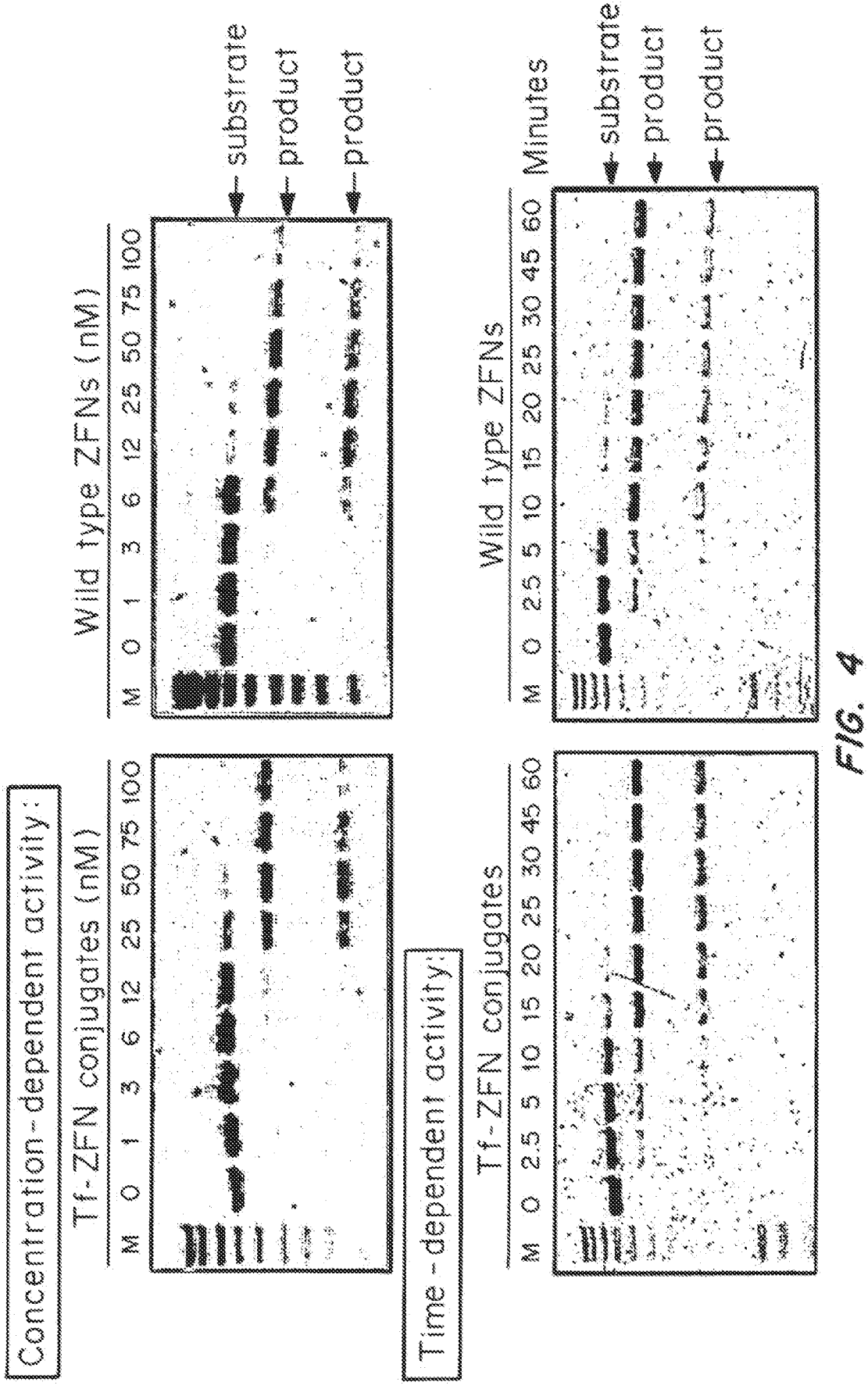

4/13
Concentration-dependent activity:
Tf-ZFN conjugates (nM)
M 0 1 3 6 12 25 50 75 100
Wild type ZFNs (nM)
M 0 1 3 6 12 25 50 75 100
substrate
product
product
Time-dependent activity:
Tf-ZFN conjugates
M 0 2.5 5 10 15 20 25 30 45 60
Wild type ZFNs
M 0 2.5 5 10 15 20 25 30 45 60
Minutes
substrate
product
product

FIG. 4

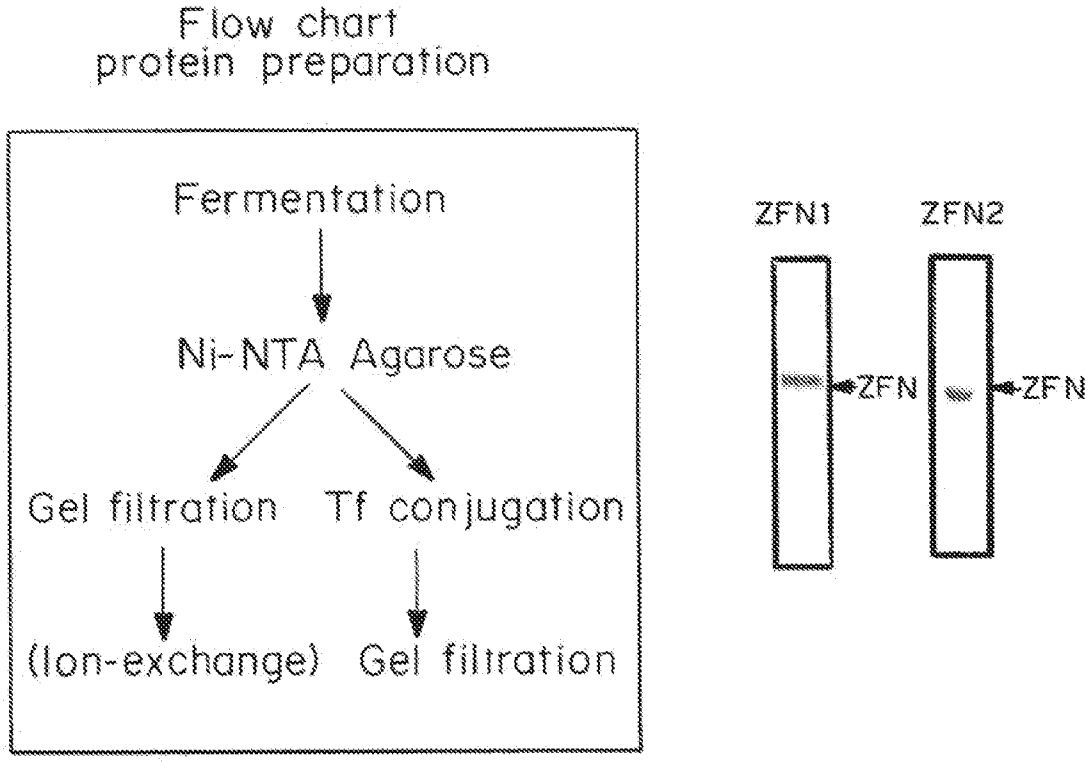

Flow chart
protein preparation
Fermentation
Ni-NTA Agarose
Gel filtration
Tf conjugation
(Ion-exchange)
Gel filtration
ZFN1
ZFN2
ZFN
ZFN

FIG. 5A

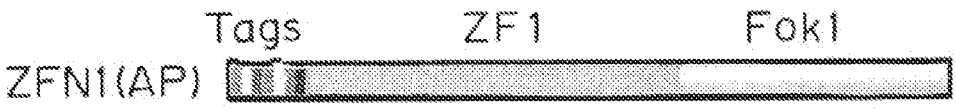

Tags
ZF1
Fok1
ZFN1(AP)

FIG. 5B

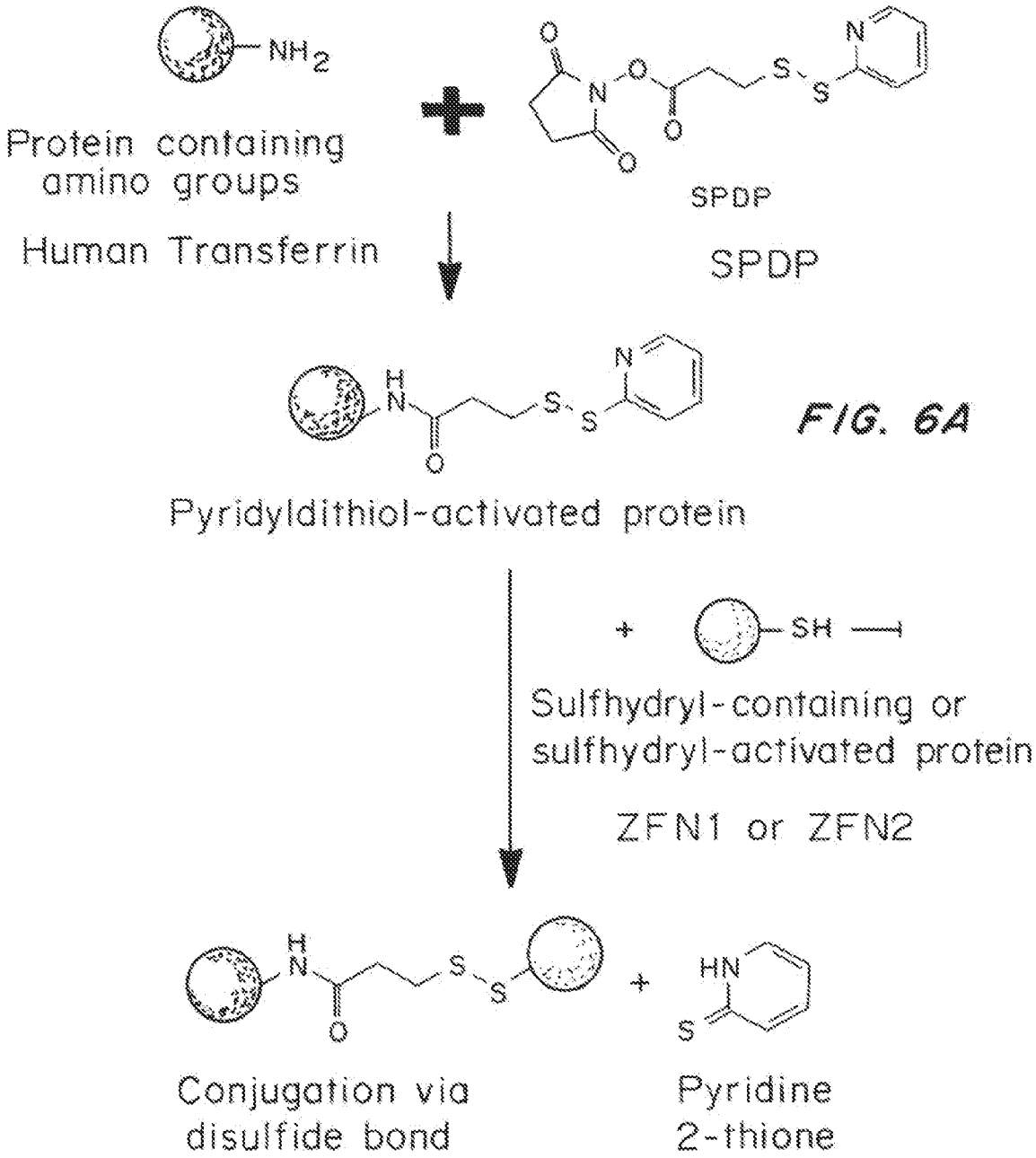

NH2
Protein containing amino groups
Human Transferrin
SPDP
SPDP
Pyridyldithiol-activated protein
SH
Sulfhydryl-containing or sulfhydryl-activated protein
ZFN1 or ZFN2
Conjugation via disulfide bond
Pyridine 2-thione

FIG. 6A

FIG. 6B

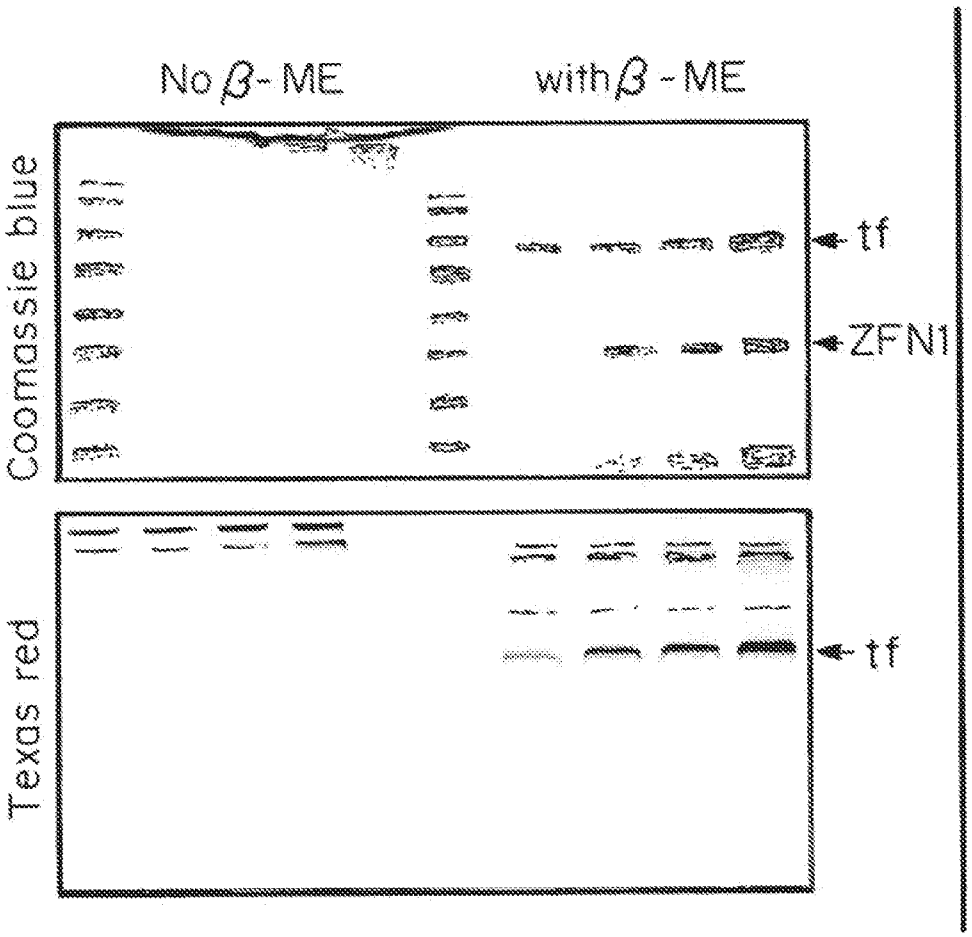


FIG. 6C

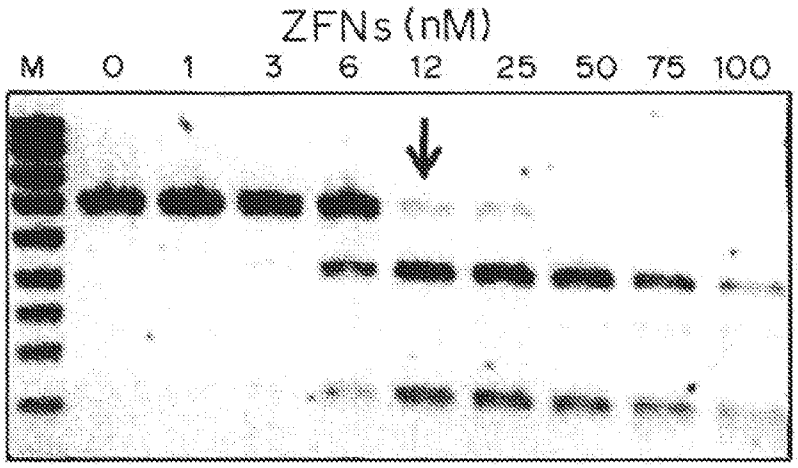

ZFNs (nM)
M 0 1 3 6 12 25 50 75 100

FIG. 7A

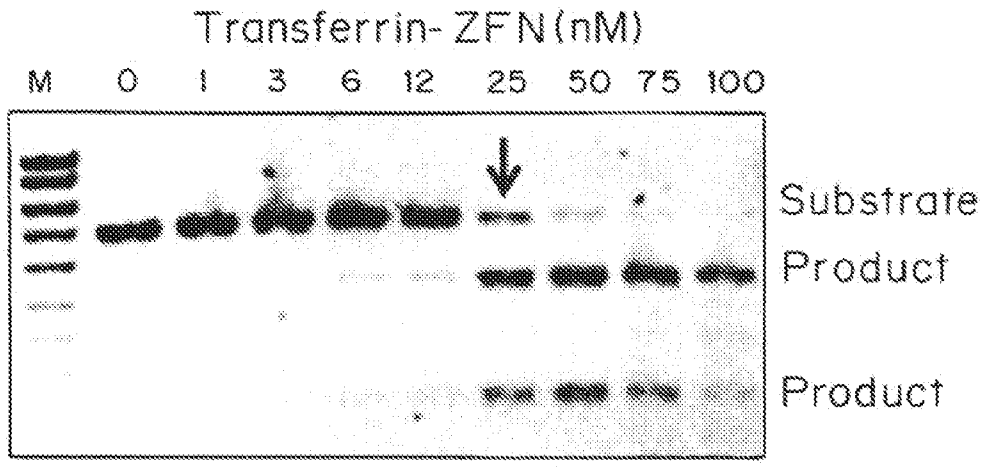

Transferrin-ZFN (nM)
M 0 1 3 6 12 25 50 75 100
Substrate
Product
Product

FIG. 7B

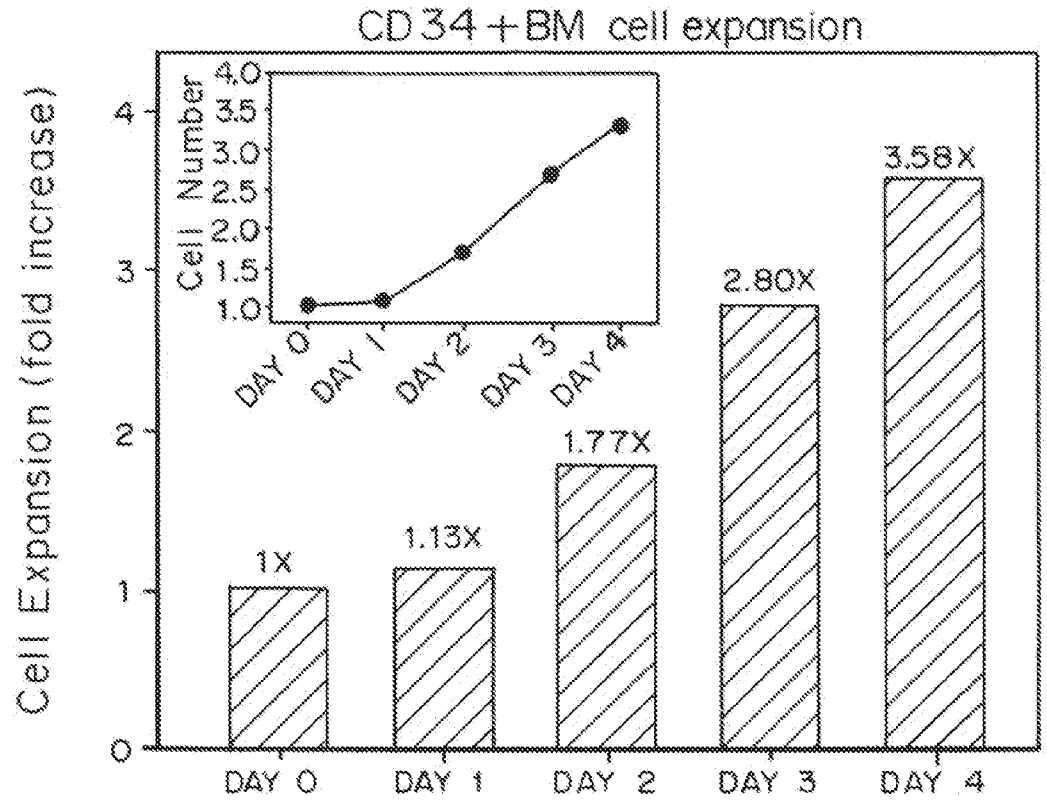

CD34+BM cell expansion
Cell Expansion (fold increase)
Cell Number
DAY 0
DAY 1
DAY 2
DAY 3
DAY 4
1X
1.13X
1.77X
2.80X
3.58X

FIG. 8A

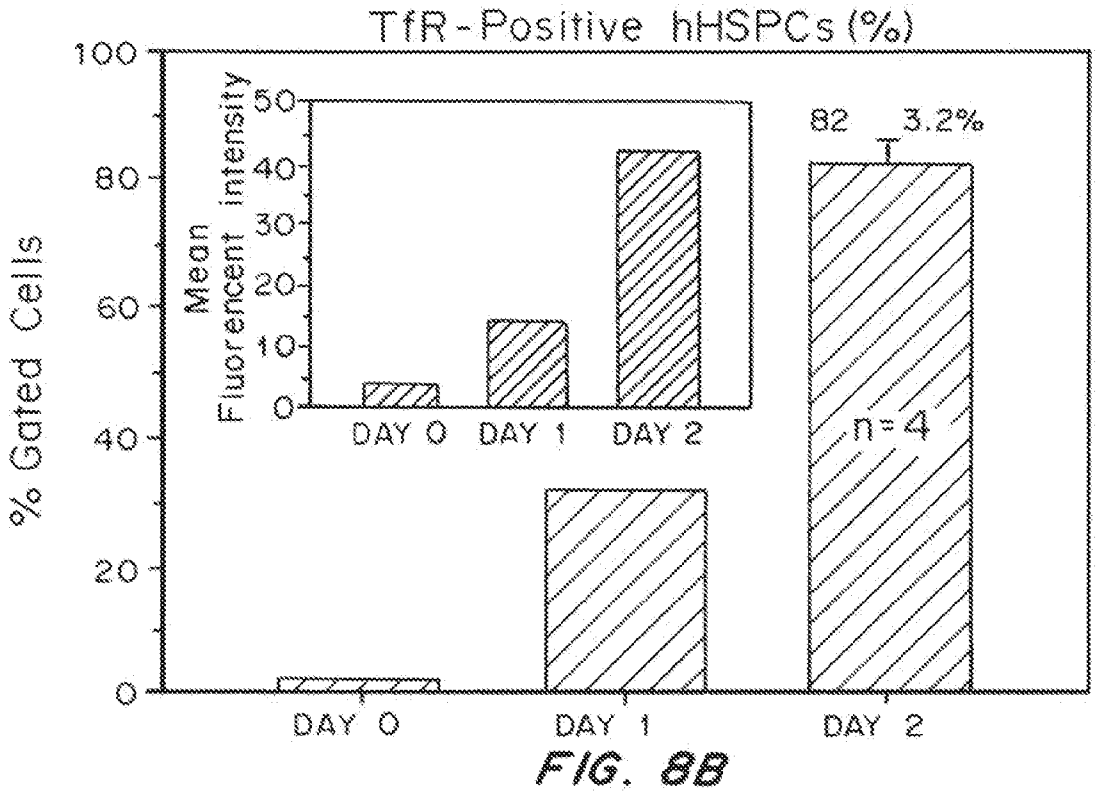

TfR-Positive hHSPCs (%)
% Gated Cells
Mean Fluorescent intensity
82
3.2%
n=4
DAY 0
DAY 1
DAY 2

FIG. 8B

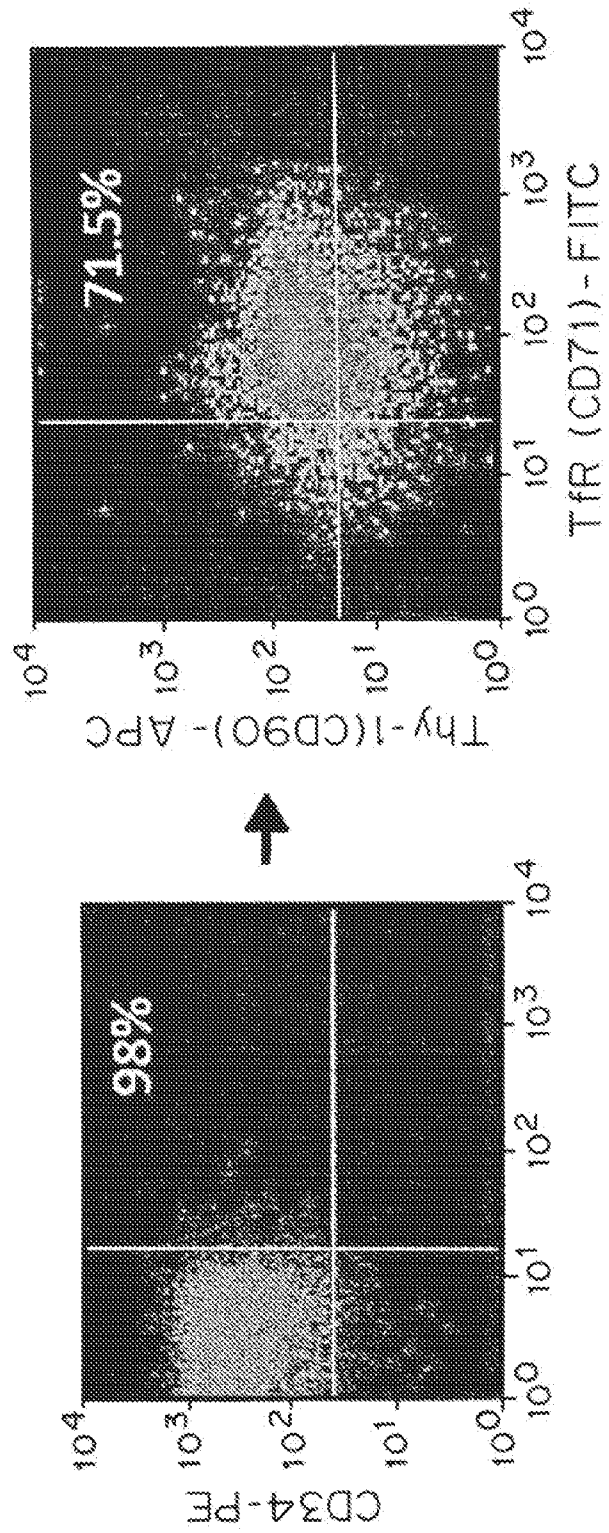


FIG. 8C

10/13

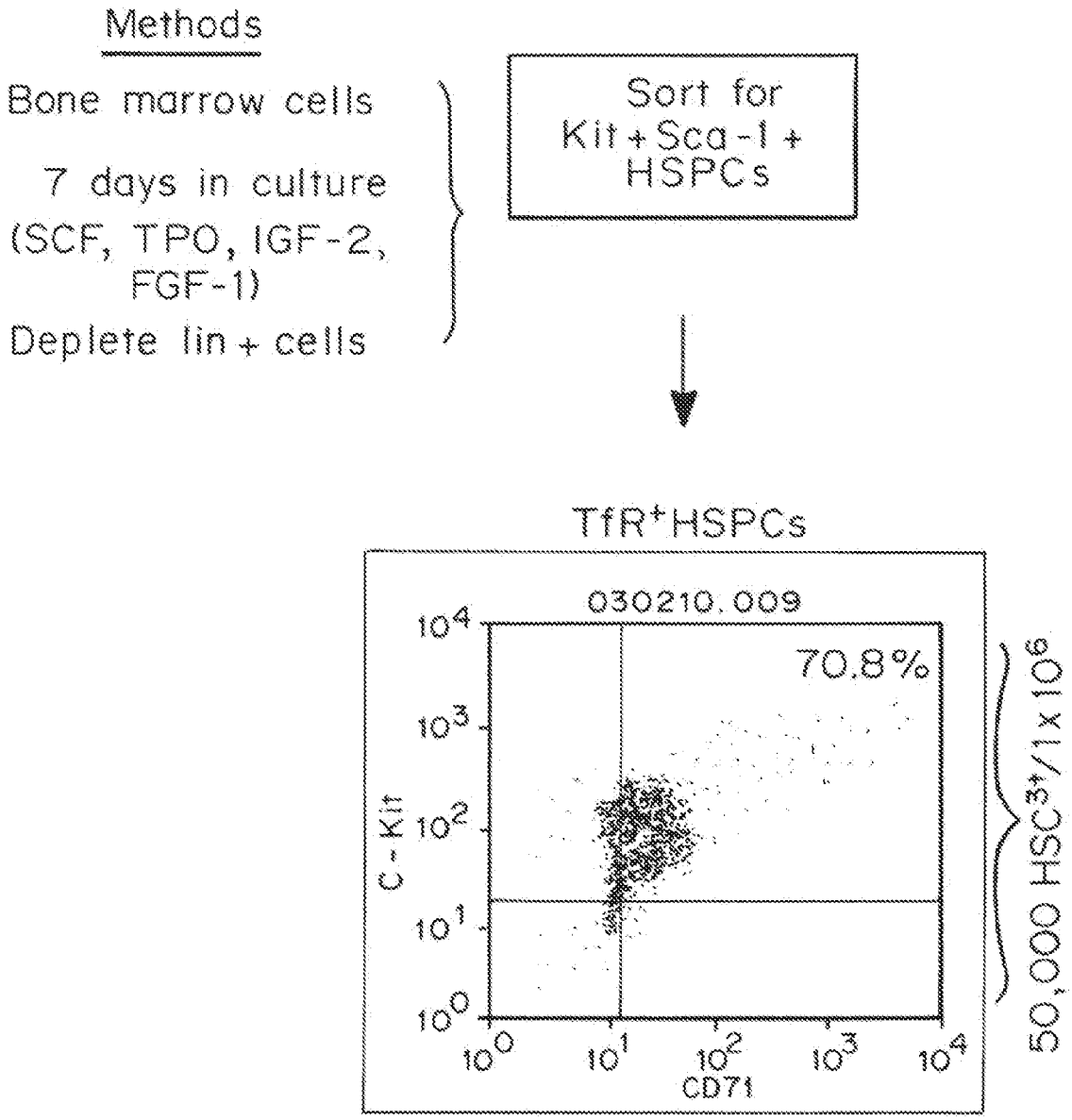

Methods
Bone marrow cells
7 days in culture
(SCF, TPO, IGF-2, FGF-1)
Deplete lin + cells
Sort for Kit + Sca-1 + HSPCs
TfR+HSPCs
030210.009
70.8%
C-Kit
CD71
50,000 HSC3+/1x10^6

FIG. 9

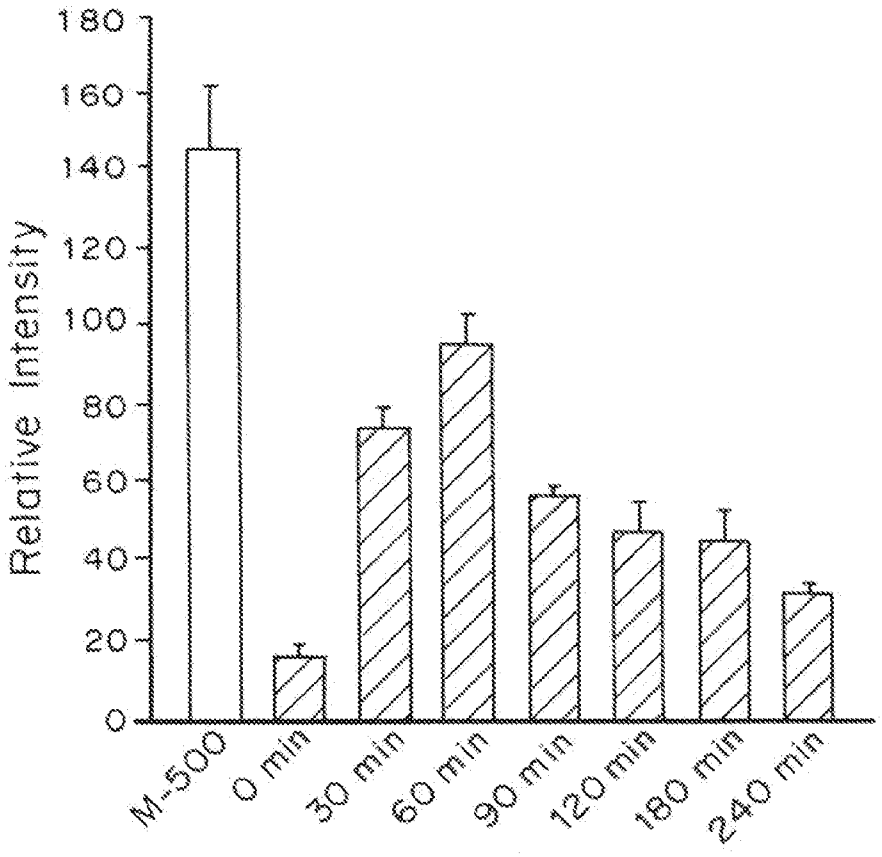


FIG. 10A

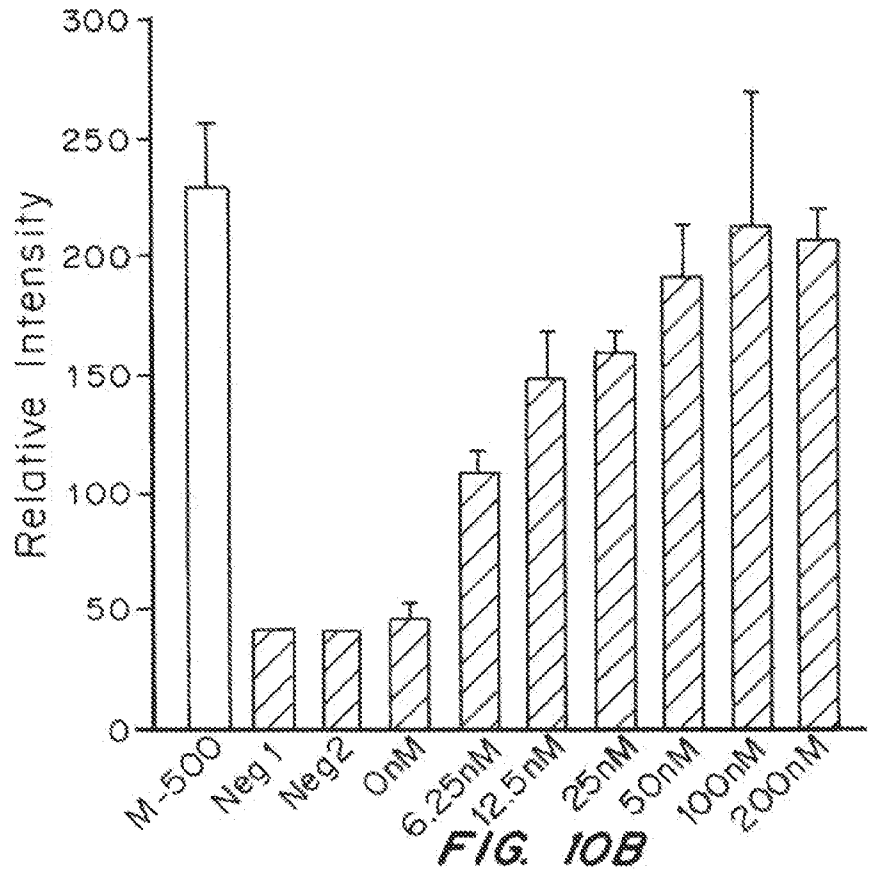


FIG. 10B

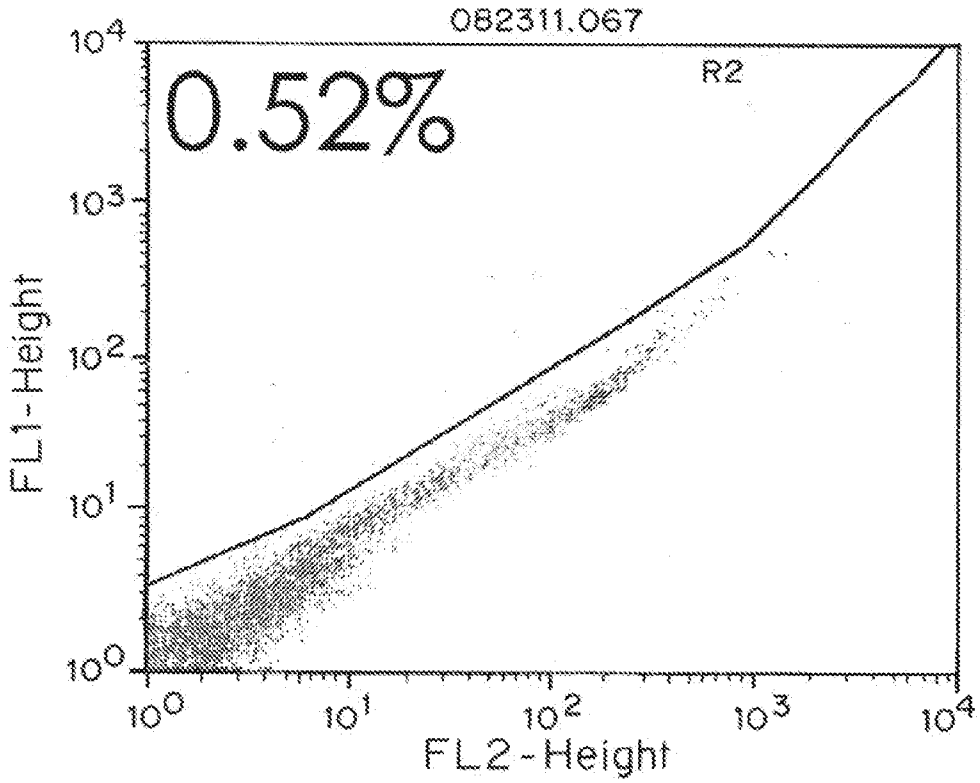


FIG. 11A

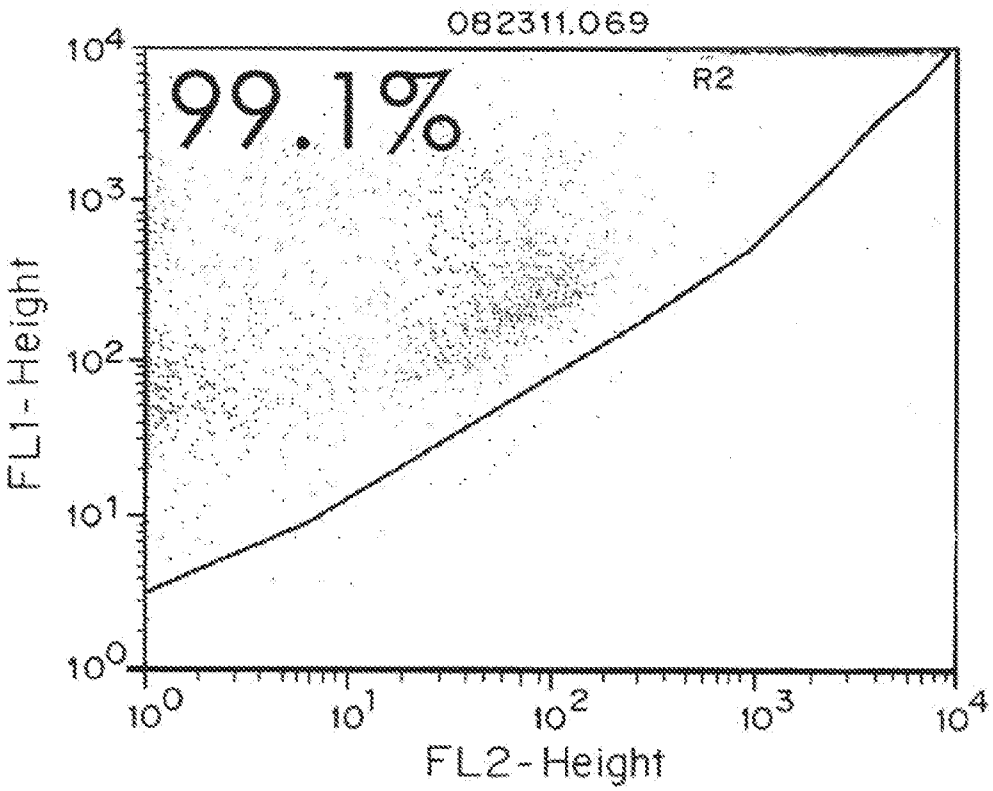


FIG. 11B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/48884

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61K 38/00; 31/7088 (2012.01)
USPC - 514/44A; 514/13.5
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 514/44A; 514/13.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 514/44A; 514/13.5
(keyword limited; terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase; Google; PubMed
Search terms: sickle cell, genome editing, zinc finger nuclease, transferring, SDF-1, homologous recombination, TAL effector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011/0030076 A1 (URNOV et al.) 3 February 2011 (03.02.2011) para [0004], [0009], [0016], [0038]-[0042], [0064], [0187], [0233], [0279]-[0280], [0282], [0361], [0372], [0379], [0383]	1-6, 10
Y	US 2011/0171179 A1 (BROYLES) 14 July 2011 (14.07.2011) abstract; para [0098]-[0099]	1-6, 10
Y	US 2011/0145940 A1 (VOYTAS et al.) 16 June 2011 (16.06.2011) abstract; para [0007]-[0008], [0202]	6, 10

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 December 2012 (27.12.2012)	03 JAN 2013
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/48884

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please see extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-6 and 10.

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Continuation of Box III: Lack of Unity of Invention

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I: Claims 1-6 and 10, directed to a method for treating a subject having sickle cell disease comprising:
a) obtaining hematopoietic stem cells from the subject;
b) delivering template DNA into the cytoplasm of the hematopoietic stem cells while in culture;
c) contacting the hematopoietic stem cells containing the template DNA while in culture with a fusion protein conjugated to transferrin or stromal cell-derived factor 1, wherein the fusion protein comprises a DNA binding domain specific for the beta globin gene and a nuclease domain that induces a double-stranded break or nick in the beta globin gene of the hematopoietic cells and promotes homologous recombination with the template DNA to replace the A to T nucleotide substitution in the beta globin gene of the hematopoietic stem cells;
d) selecting the cultured hematopoietic cells that have internalized the fusion protein conjugate and undergone homologous recombination with the template DNA; and
e) transplanting the hematopoietic cells that have undergone homologous recombination with the template DNA into the subject having sickle cell disease, wherein the transplanted hematopoietic cells produce healthy red blood cells in the subject.

Group II: claims 7-9, directed to a method of treating a genetic disorder in a subject, comprising
a) culturing stem cells or progenitor cells under conditions that increase expression of transferrin receptors;
b) delivering to the stem cells or progenitor cells template DNA comprising wild-type DNA that can replace genomic DNA of stem cells or progenitor cells comprising a mutation resulting in the genetic disorder;
c) contacting the stem cells or progenitor cells containing the template DNA with a zinc finger nuclease conjugated to transferrin, wherein the stem cells or progenitor cell containing the template DNA internalize the zinc finger nuclease conjugated to transferrin; and
d) selecting the stem cells or progenitor cells that have undergone homologous recombination with the template DNA and administering these stem cells or progenitor cells to the subject.

Group III: claims 11-15, directed to a fusion protein conjugate comprising a DNA binding domain and a Type IIS nuclease domain wherein the fusion protein is releasably conjugated to transferrin or stromal cell-derived factor 1.

The inventions listed as Groups I - III do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of the claims of Groups I-III are indicated in the Group descriptions, above.

The only common technical element shared by the above groups is that they are related to fusions comprising a nuclease, optionally Type IIS nuclease domain, and transferrin. Groups I and II are related to the treatment of a genetic disorder comprising: b) delivering to the stem cells or progenitor cells template DNA comprising wild-type DNA that can replace genomic DNA of stem cells or progenitor cells comprising a mutation resulting in the genetic disorder; c) contacting the stem cells or progenitor cells containing the template DNA with a zinc finger nuclease conjugated to transferrin, wherein the stem cells or progenitor cell containing the template DNA internalize the zinc finger nuclease conjugated to transferrin; and d) selecting the stem cells or progenitor cells that have undergone homologous recombination with the template DNA and administering these stem cells or progenitor cells to the subject. These common technical elements do not represent an improvement over the combined prior art of US 2011/0030076 A1 to Urnov et al. (hereinafter "Urnov") in view of US 2011/0171179 A9 (Broyles), as follows:

Urnov discloses a method for treating a subject having sickle cell disease (disclosed methods for targeted recombination can be used to replace any genomic sequence with a homologous, non-identical sequence; para [0279], Exemplary genetic diseases include...sickle cell anemia; para [0280]) comprising:
a) obtaining hematopoietic stem cells from the subject (such as cells explanted from an individual patient (e.g., lymphocytes, bone marrow aspirates, tissue biopsy) or universal donor hematopoietic stem cells, followed by reimplantation of the cells into a patient, usually after selection for cells which have incorporated the vector; para [0233]);
b) delivering template DNA into the cytoplasm of the hematopoietic stem cells while in culture (a method for replacement of a region of interest in cellular chromatin (e.g., a genomic sequence) with a first nucleotide sequence comprises...contacting a cell with...a first polynucleotide comprising the first nucleotide sequence (para [0038]-[0039] introduction of the replacement (or donor) sequence. The donor sequence can be introduced into the cell prior to, concurrently with, or subsequent to, expression of the fusion protein(s); para [0187]);
c) contacting the hematopoietic stem cells containing the template DNA while in culture (para [0282]) with a fusion protein, wherein the fusion protein comprises a DNA binding domain specific for the beta globin gene and a nuclease domain that induces a double-stranded break or nick in the beta globin gene of the hematopoietic cells (para [0040]-[0042]) and promotes homologous recombination with the template DNA to replace the A to T nucleotide substitution in the beta globin gene of the hematopoietic stem cells (para [0064]);
d) selecting the cultured hematopoietic cells that have internalized the fusion protein conjugate and undergone homologous recombination with the template DNA (selection for cells which have incorporated the vector; para [0233]); and
e) transplanting the hematopoietic cells that have undergone homologous recombination with the template DNA into the subject having sickle cell disease (reimplantation of the cells into a patient; para [0233]), wherein the transplanted hematopoietic cells produce healthy red blood cells in the subject (alteration of a genomic sequence in a pluripotent cell (e.g., a hematopoietic stem cell) is desired. Methods for mobilization, enrichment and culture of hematopoietic stem cells are known in the art...Treated stem cells can be returned to a patient for treatment of various diseases including, but not limited to, SCID and sickle-cell anemia; para [0282], convert the endogenous genomic copy of this mutant nucleotide pair to the wild-type sequence in a stable fashion and produce normal .beta.-globin; para [0004]). Urnov further teaches wherein the nuclease is a Type IIS nuclease (para [0045]).

- Please see next extra sheet for continuation -

Continuation of Box III: Lack of Unity of Invention

Although Urnov does not expressly recite wherein the fusion protein comprises an endonuclease conjugated to transferrin or stromal cell -derived factor 1, in a related disclosure, Broyles discloses a method of modulating .beta.-globin production (abstract), comprising administering to hematpoietic stem cells (para [0098]) a protein comprising a DNA-binding domain (ferritin-H described herein is distinct from other known trans-acting proteins in its physical properties and its proposed function as a repressor that binds primarily to CAGTGC [SEQ ID NO: 6] sequences in the .beta.-globin and .gamma.-globin promoters; para [0099]) and transferrin (para [0098]) for the purpose of targeting receptors on the surface of the cell (para [0098]) to improve uptake of the fusion protein.

It would have been obvious to a person of ordinary skill in the art to include a tranferrin fusion domain, as taught by Broyles, to a dna-binding fusion protein, as taught by Urnov, in order to improve the targeting and uptake of the protein by the target cells, as taught by Broyles, when providing the purified protein, itself, to the cells, as taught by Broyles, as opposed to expressing the fusion protein within the cells, as taught by Urnov.

Therefore, the inventions of Groups I-III lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.